Cross-Language Phonetic Priming in Bilinguals

A thesis

submitted in partial fulfilment

of the requirements for the Degree

of

Master of Arts

in Linguistics in the

University of Canterbury

by

Keyi Sun

University of Canterbury

2011

Abstract

This thesis looks at cross-language phonetic priming effects on late L1-dominant bilinguals, with different degrees of proficiency within the group. The participants in the study are 14 Chinese-English late bilinguals, whose production of vowels and consonants in different priming language contexts was analysed. The 14 speakers were divided into two groups based on their language proficiency. Information collected from questionnaires in two different languages was used to divide them into the two groups. They were required to participate in the experiment in two different sessions. On one occasion the interviewer spoke English to them and this was followed by their English reading and Chinese reading; whereas on the other occasion the interviewer spoke Chinese and the subjects did the opposite reading order from the first condition. Significant results of the analyses show that non-early, L1-dominant bilinguals do not differ in proficiency across priming conditions. Both groups show significant changes as the result of language priming for exactly the same vowels and the same consonants. Significant changes in the production of the sounds reveal interference between certain L2 sounds and their L1 counterparts. However, near significant results also show an unexpected direction of changes in production in L2, which may have been caused by experimenter identity. Furthermore, transfer effects of L1 on L2 found only among high proficiency speakers suggest that inhibitory control is dependent on L2 proficiency.

Table of Contents

List of Figures

List of Tables

Acknowledgments

To the University of Canterbury Linguistics Department, thank you for providing a nice environment to study.

I would like to express sincere appreciation and gratitude to Professor Jen Hay, who assisted me with great help and encouragement, and provided me with excellent ideas. This thesis could not have been finished without your wisdom and supervision.

I would also like to thank Dr. Hong Xiao for her assistance in the field of Chinese linguistics. Your observations gave me the initial inspiration for possible testing materials. You also kept me on track during the earthquakes; thanks for the lovely dinner.

I would also like to thank Robert Fromont, who helped me create textgrid files by using the HTK Toolkit. Special thanks to my friend Joe Zhu, who was available to help me with my statistic problems. Sometimes I just needed to knock on your door.

Lastly, I would like to acknowledge all 14 speakers who dedicated their time to participate in my experiment. I know it was hard for some of you; especially since you all had to come to me twice. Thanks to Shawn for his extended social network; although I did not know all of your friends, they are nice people. Without them this thesis would not exist.

Chapter 1

Introduction

People accommodate their speech during social interactions. Bilinguals can select a particular language in response to other speakers. Some researchers have tested whether recent exposure to one language can affect the production of the other for early bilinguals. However, it remains unclear whether short term priming is correlated with language proficiency for non-early second language learners. The current study is going to address this question.

In second language acquisition it is often claimed that people cannot achieve a native accent if they begin to learn the language after around age ten (Lenneberg 1969). What is often studied is the effect of the first language (L1) on the second (L2), i.e. which features second language learners bring from their mother tongue into the second language. For example, learners often transfer the closest sound from their L1 into a sound learned in L2. Studies on the effect of L2 on L1, on the other hand, are relatively few. Some scholars, for example Pavlenko (2000) and Cook (2003), studied L2 effects on L1. Their studies suggest that the effect of L2 on L1 can be found on various levels such as lexical, semantics, phonological and pragmatics. Kecskes and Papp (2000) found that students who took a special course on a foreign language did better in written language in their native language. Su (2001, 2010) found that learners of English whose mother tongue is Chinese carried L2 strategies into L1 processing; and a bi-directional effect was also found on a pragmatic level. One might ask if the influence of L2 on L1 can occur at the level of phonetics and phonology, and the answer is positive. Sancier and Fowler's (1997) study suggests that speakers' production of a particular sound can shift toward the ambient language environment; this includes shifts in L1 when the ambient language is L2. In an early experiment conducted by Flege (1987), he found that French production by French-dominant bilingual speakers was also affected by their English. Flege argued that the age of arrival (AOA) (1992a, 2002), length of residence (1988) and L2 speaking targets (2002) all play important roles in a bilingual speaker's L1 accent. Major (1992) also tested the hypothesis that L2 proficiency is correlated with L1 loss by comparing bilingual speakers' production with two monolingual control groups, and the results confirm that such a correlation does indeed exist. The current study is going to examine the effect of short term exposure to an L2 environment; specifically, it is going to look at the effect of short-term priming of one language context on another. It should

be pointed out that the 'recent exposure' in current study is very short. Sancier and Fowler (1997) conducted a research on the effect of recent language experience of six months on the other language. Recent exposure in the current study is only within the experimental environment.

The focus of the present study is the effect of L2 on L1 in the area of phonetics and phonology. It tests the bidirectional effects of speech context among Chinese-English bilingual speakers according to their level of proficiency in English. Age in learning L2, age of arrival (AOA), length of residence and L2 speaking input are used for grouping our participants. Individual information was collected in two questionnaires. One was designed to collect information on participants' personal background, and the other to collect information on their daily language use. We then tested each participant in two different language environments, in order to observe their speech production. We were interested to see how priming with a particular language affects the other, and if such effects are correlated with language proficiency.

Finally, we used statistical analyses to compare the results from these bilingual speakers with different L2 proficiencies. We were therefore able to identify the influence of the level of L2 proficiency on the degree of change in L1, and to answer our research question concerning the effect of recent short-term exposure to L1 and L2.

The overall structure of this thesis is as follows:

The thesis begins with an introduction of existing studies and literature in the area of language influences between L1 and L2. Chapter 2 reviews findings from existing studies and important theories in Second Language Acquisition (SLA) and priming effects. On the basis of this discussion of reported effects of L1 on L2 and L2 on L1, as well as studies about short-term priming, we then present our own research questions.

Chapter 3 presents the methodology, including the procedures of the experiment, the recruitment of the participants and the way the classification of L2 proficiency was carried out. It also describes the research material, e.g. the sounds, including comparisons

between English and Chinese, sound selection and reasons for choosing the sounds at the centre of this study.

Chapter 4 presents the results for each vowel. Segments are presented separately and so are speakers of both proficiency levels for each segment or pair of segments.

Chapter 5 presents the results for each pair of consonants. Segments are presented separately and so are speakers of both proficiency levels for each pair of segments.

Chapter 6 discusses the results and implications and considers problems and unexpected results.

Chapter 7 provides an overall summary of the thesis, highlights important findings, and makes suggestions for future research.

Chapter 2

Background

The purpose of this chapter is to provide background information for the current study on the basis of existing literature. It discusses relevant findings and outlines the motivations for the current research. To identify relevant findings, we need to first distinguish Second Language Acquisition (SLA) from bilingualism. SLA usually refers to a learning process of a language other than one's first language, and also tends to be reserved for people who acquire the second language much later than the first; bilingualism is the result of second language learning. Bilingualism is also the aim of SLA. However, in practical use, since bilingualism can have different degrees, what is often called 'bilingualism' is often used to refer to various stages of second language learning or using: in Flege (1987, 1988, 1995 and 2002) the word 'bilinguals' is widely used instead of 'second language learners'. In this thesis, bilingualism refers to the possible various stages of SLA, whereas SLA refers to the process of second language learning.

2.1: Important models of bilingual representation

Researchers have proposed different types of models for bilingual representation. For example, Lenneberg's (1969) research suggested that people's ability to learn L1 disappears after age 12, which also raised the question of learning L2. However, this Critical Period Hypothesis cannot explain the success of L2 learning for some adults. Some other theories, such as Contrastive Analysis, argue that if an L2 sound is similar to an L1 one, learning the sound would be easy, and vice versa. However, when Flege (1987) was conducting a research on sound production between different French and English vowels, results showed that a French vowel, which was new to English speakers, was produced significantly better than another French vowel which had a counterpart in English. This was contradictory to Contrastive Analysis.

When talking about the relationship between two languages in one's mind, Cummins (1984) suggested two different models for bilingualism, the Separate Underlying Proficiency (SUP) and Common Underlying Proficiency (CUP). The SUP states that there is a balance between L1 and L2. When the proficiency in one language increases, the proficiency in the other language would decrease. It assumes that human brain cannot accommodate two full languages, and there is no interference between the two language systems. The SUP failed to explain early bilinguals who

have proficiency in both languages; because it assumed that it is impossible that human brain can hold two fully acquired languages. On the contrary, the CUP assumes a common space shared by both languages. The development of either language would cause the development of proficiency in both languages. However, if learning L2 can also promote L1 proficiency, we might need to ask why people report loss of the first language by immigrants as the second type of language loss (Kouritzin, 1999). The idea of 'common space' was also argued for by Cook (1992, 2003). However, in contrast to the CUP, Cook assumed that the relationship between the properties of the two languages is not fixed, but changing. There would be interference between elements in the shared space, but not those that are not in the common space.

An important framework in recent years was suggested by Flege (1987, 1988, 1995; Flege, Schirru & MacKay, 2003) as the Speech Learning Model (SLM). The SLM proposes that L2 learning is largely dependent on L2 input, and after certain length of time of input, L2 learners can form knowledge of L2 sounds. The SLM also suggests that the ability to learn L1 does not disappear through aging, thus it predicts that even adult learners can achieve native-like L2. According to the SLM, the phonetic properties of L1 and L2 share a common space, so the two phonetic systems can affect each other. The SLM argues for an effect of similarity. It involves the occurrence of category formation when an L2 sound is perceived as 'new' or 'similar'. It states that a possible reason for the failure of some bilingual speakers to acquire a particular sound in L2 is the similarity between a sound in L1 and the target sound in L2. These speakers do not treat a L2 sound as a 'new' sound but a 'similar' sound. As a result, instead of forming a new category for the new sound, L2 learners combine the new sound from L2 into their existing category of the corresponding sound from L1. The SLM predicts that category formation becomes less likely, but not impossible for children through aging for an L2 sound which is close to an L1 sound. This can well explain some foreign accents in SLA. The SLM also predicts that a sound will dissimilate from a neighbouring sound when a new category is founded for the sound. Brown (1998) also pointed out that properties of speakers' L1 can be translated into L2, consistent with Flege's hypothesis, but what determines the translation is the speakers' perception of L2 properties. Kang and Guion (2006) found a merged category for late Korean-English bilinguals.

In their study, they compared sound production of English and Korean stops from both early and late bilinguals. They found that early bilinguals produce more native-like stops of both English and Korean; in contrast, late bilinguals produce English stops that are similar to Korean stops.

It is also pointed out by some people that L2-dominant bilinguals are less likely to have such inter-lingual interference. Flege, MacKay and Piske (2002) found such a result from their experiment by comparing the production of English sentences produced by early and late Italian-English bilingual speakers who also exhibit a difference in L2 dominance. In their study, Italian-English bilinguals were recruited and divided into four groups basing on their different ages of arrival and amount of L1 use. As a result, they had four groups of early-low, early-high, late-low and late-high speakers. Both early and late bilinguals produced English sentences with an accent but another group of early bilingual who were English-dominant selected from the subjects did not produce English sentences with accent. Lambert (1990) also found that balanced bilinguals (using L1 and L2 at nearly equal dominance) show less interference between L1 and L2. If Flege, MacKay and Piske' finding is true, then L1-dominance bilinguals would be more likely to have interference between the two languages.

2.2: Transfer Effects

Effects of L1 on L2

As has been discussed, the effects of L1 on L2 have been modeled by many researchers in approaches such as the SLM. Flege (1981) coined the term 'phonological translation'. He argued that 'phonological translation' is based on a speaker's existing L1 and the pairs of corresponding sounds found in L1 and L2. For example, Zampini (1994) mentioned that English native speakers have difficulty producing Spanish /b d g/. English /b d g/ never undergo spirantization, so these sounds are too often produced as stops by English learners of Spanish, which make English speakers sound strange when producing Spanish /b d g/ without the frication they require.

Effects of L2 on L1

Several studies have revealed that an effect of L2 on L1 can also be found on various levels of language acquisition. An assumption has been made as early as in Flege's study (1981), where he mentioned a prediction that the amount of use and exposure to L2 would eventually cause sounds in L1 to undergo change. Later when Flege (1987) introduced the Speech Learning Model (SLM), his experiment showed that bilingual speakers' L1 production of plosives was affected by their L2. His research results showed that French-English bilingual speakers' production of French plosives is somewhere intermediate between native French plosives and native English plosives. A similar finding reported by Major (1992) also suggests that the degree of mastering one language has a positive correlation to the loss of the other language.

2.3: Inter-lingual mutual effect between L1 and L2 on Phonetics

Sancier and Fowler's (1997) study suggests that a bilingual speaker's L1 or L2 production of a particular sound can shift toward the ambient L2 or L1 language environment, thus causing the L1 or L2 to undergo change. In their research, they tested a bilingual speaker's production of plosives. Their findings raised challenges to what is known as 'the critical period'. The subject of Sancier and Fowler's study had passed puberty and was living in the US, but she frequently travelled back to Brazil. Her production of plosives shifted in both languages, more English-like or more Brazilian Portuguese-like every time when she travelled to either of the destinations after spending some time in the other. This particular case suggests that the two co-existing languages in one speaker's mind can influence each other. Sancier and Fowler's study looked at only one speaker, whereas in the current study, analyses are conducted across a number of speakers.

2.4: Factors affecting competency in L2

Previous studies suggest that certain factors can affect the overall success in SLA. These factors are discussed below.

Age of Arrival

Age of arrival (AOA) is an important factor. Age of arrival has been studied in L2 production of consonants (Flege et al., 1995; MacKay et al., 2001), and vowels (Flege, 1992b; Flege et al., 1999a; Piske et al., 2002). It is widely believed that the age of first exposure to a L2 can influence the overall performance in the L2 (Flege, 1992a; Flege, et al., 1999b; 2002; Piske et al., 2002). People who arrived in a country in which the bilingual's L2 is the native language earlier in their lives perform better than those who arrived in adulthood. One study carried out by Flege et al., (1995) looked at some Italian subjects who arrived in Canada at different ages. The results show that there is a positive correlation between AOA and the degree of accent: the older the subjects were when they arrived in Canada, the stronger their accent. Another study also highlighted the importance of AOA: L2 consonant identification can only be achieved as accurately as that of speakers who use L2 as a native language when L2 learning starts early in life (MacKay et al., 2001). AOA is also important for vowel acquisition. Munro et al. (1996) found that native Italian speakers' accuracy in producing English vowels was correlated with AOA. Similar results were also found by Flege et al. (1999a) and Piske et al. (2002), who showed that the accuracy of producing L2 vowels correlated with children's age of arrival.

L1 use

Another factor mentioned by Flege is L1 use. In an experiment (Flege, 2002) Flege showed that not just AOA, but the degree of L1 use also influences bilinguals' production of L2. Piske et al. (2002) also demonstrated that the accuracy of L2 vowel production is influenced by the amount of L1 use. By comparing L2 vowel productions between early bilinguals with high usage of L1 and those with low usage of L1, Piske et al. found that the more bilingual speakers continue to speak L1, the less accurate L2 vowels they produce.

L2 input and length of residence

The third important factor is length of L2 input and who L2 learners speak with. Flege and Liu (2001) examined Chinese-English bilinguals living in the US with different lengths of residence. The participants consisted of students and non-students. The results revealed that among students, those who had experienced longer residence achieved better scores than those who had lived in the US for a shorter time. As for the scores of the subjects who were non-students, there was no significant difference among them in terms of length of residence. Interestingly, students and non-students were not different in terms of the percentage of L2 use according to their self reports. The study highlighted the importance of native input in SLA. Despite the fact that frequency of L2 use is important; whom L2 learners speak with heavily influences the result.

2.5: Contextual Effects

There are researchers, who claim that the production of either language by a bilingual speaker depends on what language mode they use. In other words, people accommodate during speech to gain 'social approval' from their listeners (Coupland, 1984). It was pointed out by Grosjean (2001) that bilingual speakers' speech productions are different when they speak to different monolingual speakers. For bilinguals, the language mode can be either monolingual or bilingual, and speaking to monolingual speakers will trigger the corresponding language mode. In the current study, the experimenter is a Chinese-English bilingual, who will create English context for other Chinese-English bilinguals. It will be interesting to see whether their language mode is triggered by the context or by the experimenter identity.

For different language modes, a so-called 'situational continuum' states that the production of a bilingual largely depends on the different levels of activation of both languages. Apart from the motivation for social reasons among adults, some other researchers showed the effect of priming of foreign-language input could also affect infants (Kuhl, Tsao & Liu, 2003). In their research, they expose nine month old American infants to native Mandarin Chinese speakers, while a control group was only exposed to recorded Chinese. The results show that the differentiation between native and foreign sounds could only be achieved by participating in social interaction.

The effect of experimental material and environment on speech production and perception can be controlled for giving experimental material in only one language (Grosjean & Miller, 1994; Hazan & Boulakia, 1993; Sundara, Polka & Baum, 2006). Antoniou et al. (2010) argued for the importance of experimental material and revealed that by giving material in one language, bilingual speakers would activate only one language to fit into the language context and inhibit the other one. In the research conducted by Antoniou et al., they tested early Greek-English bilinguals' production of both Greek and English stops in only one language context at a time. Thus the contexts could guarantee the language mode. They found that in the Greek context, plosives of Greek were produced as same as produced by Greek monolingual speakers; whereas in the English context, the bilinguals' production of English plosives at word-initial position was similar to the production of English monolingual speakers. In the current study, we will also use different contexts to trigger language modes. We are particularly interested in seeing whether priming effects are different across different proficiency groups.

2.6: Research questions and the contribution of the present study

The present study will examine the following issues.

1. Effects of short term exposure of L1 on L2.
2. Effects of short term exposure of L2 on L1.
3. Investigating whether effects of short-term exposure are dependent on L2 language proficiency

Researchers have found the influences of the two languages on each other in bilinguals. Both Flege's SLM (1995) and Cook (1992, 2003) assume there is a common space in a bilingual's mind for elements from both L1 and L2 to share. However, as some others have pointed out (Lambert, 1990; Flege, MacKay & Piske, 2002; Kang & Guion, 2006), for early and L2-dominant speakers, they appears to be more independence between L1 and L2 systems. Logically, late and L1-dominant bilinguals would be more likely to exhibit dependence between the systems. In the current research, all the speakers are non-early, L1 dominant bilinguals. They are divided

into two groups according to their English proficiency. We are going to see if the two groups behave differently when exposed to different contexts. If priming with one language causes the other language to change in the language used by the individual participants, this would show that there is interference between the phonetic properties of the two languages.

Chapter 3

Methodology of the Production Experiment

The experiment was designed to reveal contextual effects on the production of L1 and L2. Thus the experiment is based on an assumption that speakers will only activate one language mode at a time. To see the effect of context on both L1 and L2, both languages need to be spoken in context of the other language.

3.1 Procedures

The experiment involved filling in questionnaires (see Appendix D) and reading passages (see Appendix B) in English and Chinese. We required that each participant needs to participate in the experiment in two different sessions, and the two sessions were at least one week apart. The two sessions differed in terms of linguistic context and order of two reading passages. Diagram 3.1 illustrates the concept.

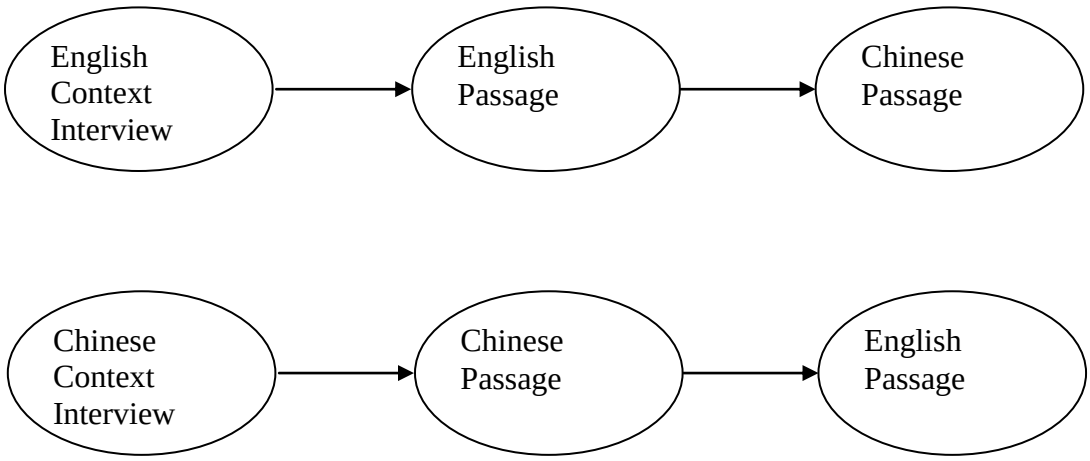

Figure 3. 1: Speech contexts and order of the passages

In one of the sessions the interviewer spoke English and in the other the interviewer spoke Chinese. In the first session before the interviewer began to go through a questionnaire, each participant needed to sign a consent form (see Appendix F). The consent form only needed to be signed once. In each session, the participants filled in the questionnaire directly prior to reading the passages. To create a better language context, the interviewer went through each question in the questionnaire with each participant. One questionnaire was in English and directly followed

by the reading of the English passage and the other one was in Chinese, and directly followed by the reading of the Chinese passage. Both passages were read in each session. Since the English passage was read right after the introduction in English, the priming in English would guarantee that the English production is English-primed. In addition, the English passage reading also strengthened priming for the following Chinese passage reading, which was English-context-primed. The same went for the Chinese context interview and its following Chinese and then English passage readings, both of which were Chinese-context-primed. All the readings for all the participants were recorded in a same quiet room, on the University of Canterbury campus. In the second session, after reading the passages, the participants were told the aim of the experiment by being presented with a debriefing sheet (see Appendix G). The study aimed to compare the pronunciation across the language contexts, between high proficiency English speakers and lower proficiency speakers. About half of the participants read the Chinese passage in the first session whereas the other half read the English passage first to counter balance. The counter balance was also equal within high and low proficiency groups. Of seven speakers who read English first, four were high proficiency speakers and three were low proficiency speakers. Of seven speakers who read Chinese first, four were high proficiency speakers and three were low proficiency speakers. The consent form also informed the participants that the experiment required them to come twice and a stated aim of the experiment was to investigate pronunciation patterns in bilingualism. No-one questioned why they needed to read passages in the reverse order for the second time. Some of them questioned why the experimenter spoke different languages on the two occasions, and they were told that it was a requirement of the experiment. The equipment used in the experiment was a digital recorder TASCAM HD-P2, with a 4GB Lexar CompactFlash card and a noise-filtering AKG MPAIII C444 head-worn microphone. The settings used were 16 bit resolution at 44.1 kHz sampling rate and mono channel.

The experimenter was a late Chinese-English bilingual speaker, age 29, who has spent seven years in New Zealand. Since the participants were mostly recruited with public notices, emails and through extended personal networks, most of them did not know the experimenter. The language they tended to originally use to respond to the notices and emails was English. The

notices and emails were all in English. It has been shown that people can shift their speaking styles when talking to different community identities or interviewers (Bell, 1984; Rickford & McNair-Knox, 1994). As a result, the experimenter kept speaking English to the participants who met the experimenter for the first time and needed to do the English context first, and when they came back for the second occasion, the experimenter spoke Chinese. This was to minimize the effect of ethnic identity. The identity of the experimenter might have a possible effect on the participants. Few participants who knew the experimenter had to do the Chinese context first and at the end of the first session they were told they needed to participate in a completely English speaking environment for the second occasion. There were four participants who knew the experimenter and all of them did the Chinese context first.

3.2 Questionnaires

The two different language contexts were created by filling out two different questionnaires under the experimenter's guidance. The experimenter asked the questions in an interview format, and filled out their responses. There were two questionnaires – one for each session, and they were in different languages. In each session, the language of the first reading passage was the same as the language of the questionnaire. Therefore, the questionnaire ensured that the participant was speaking the 'priming' target language immediately prior to reading the passages.

The other aim of the questionnaires was to collect information from the participants, which was then used to classify the participants' L2 proficiency. The questionnaire reflects the consideration of the effect of age of arrival and other aspects. It collected information on age of arrival (AOA), length of residence (LOR), L2 input/frequency of contact with local people, L1 and L2 speaking frequency, and language attitude. The questions were divided into two parts. Questions in the English questionnaire were mostly related to personal information, such as age, gender, AOA, LOR, occupation, living situation and family situation; and the questions in the Chinese questionnaire were mostly related to language use and attitude, such as percentage of L1 and L2 use, L2 input and children's language use. The questions were divided across the two questionnaires to avoid a sense of repeating questions. The two sessions were at least one week

apart; and some participants came back to the second session much later than that. It is suspected that few of them could remember the content. A possible effect of the Chinese questionnaire on the following Chinese reading part will be discussed in the discussion chapter.

Based on their answers to the questionnaires, the participants were classified as high proficiency L2 speakers or lower proficiency L2 speakers. This classification is relative. The judging process involved integrated evaluations. In other words, the participants' personal statuses are all taken into consideration. Each of their answers was given a score; a high score represents relatively higher proficiency. Personal characteristics include percentage of daily talking conducted in Chinese, the language spoken by most of their friends, AOA, LOR, degree, partner and the language spoken at work.

3.3 Participants

The participants were recruited through personal extended networks as well as some public notices. Information about the study was posted on public message boards and internet forums, such as Facebook.com and bbs.skykiwi.co.nz, the latter being an internet forum that attracts a large number of Chinese people living in New Zealand. Bilingual speakers were recruited in Christchurch, and included both students and non-students. The age range of the sample is 22 to 34. The places the participants originally came from are Harbin, Hebei, Shaanxi, Shandong, Yunnan, Hubei, Guizhou and Jiangxi. Apart from Jiangxi, all belong to the North Mandarin region. Despite the fact that they come from different North Mandarin regions and North Mandarin also has dialectal variations, they all speak standard Chinese, or Putonghua, as a result of their education and Putonghua promotion in China. Each relevant answer on the questionnaires was given a score. Criteria are given below.

	score	Example
% of Chinese	(1-%)x10	(1-20%)x10=8
Friend	3 - more English, 0 - more Chinese	3
AOA	12-AOA	12-24=-12
LOR	LOR	10 years = 10
Chinese partner	3 for no, 0 for yes	3

English at work	3 for yes, 0 for no	3
degree	1-4 gradually	2 for Bachelor degree

Table 3. 1: Scores for answers to the questionnaire

Percentage of daily talking conducted in Chinese/English is more important for it reflects L2 input, and it has more weight. AOA takes age 12 as a standard point, based on the theory of the Critical Period in language acquisition. LOR has a positive effect on L2 speaking, thus it also has more weight. Both friends and the language spoken at work have possible effects on L2 input. The language spoken by most of their friends can relate to L2 input; however we did not know whether their friends spoke native English, so it was given lighter weight. Same reason for the language spoken at work, and the two were given equal weight. Having a Chinese partner partly reflects Chinese usage at non-working occasions, but no additional information was collected on the length of a working day. As a result, it was given equal weight as English usage at work, but with 3 points for not having a Chinese partner. Degrees reflect the length of time spent in education. It based on a study from Flege and Liu (2001) which found a difference in English production between bilingual students and non-students.

Taken all together, the range of possible scores is from -10 to 25. The median point of the range is 8, so lower than 8 will be categorized into the group of low proficiency speakers (see Appendix C). As a result, the fourteen speakers were divided into two groups. The results are given in Table 3.2.

	high	low	total
male	6	4	10
female	2	2	4
total	8	6	14

Table 3. 2: A summary of participants' genders and L2 proficiencies

3.4 Materials

3.4.1 Target sounds

The purpose of this section is to give background information about our testing targets, i.e. tested sounds. It discusses the characteristics of some sounds and some reasons for choosing these particular sounds. It also discusses some relevant findings from existing studies.
The present study is concerned with sound production as being affected by both L1 and L2 of a bilingual speaker. Thus, the selection of testing targets needs to fit our purposes. The criterion for selecting the testing targets is as follows.

> A pair of sounds with one from each language should be similar but not identical across the two languages.

The criterion is based on Flege's second language learning model (1987). When a L2 sound is perceived as 'similar', category formation is unlikely to occur. Most existing studies on L1-L2 effects took plosives as testing targets. In the experiment conducted by Kang and Guion (2006), they used stops as testing targets. However, in our case, plosives are not a good choice, due to the findings from Deterding and Nolan (2007). In their experiment, Deterding and Nolan (2007) recruited standard Chinese and RP British English speakers to read a list of Chinese characters and a list of English words. The lists contained six plosives, including three voiceless unaspirated plosives /p/, /t/, /k/ and three voiceless aspirated ones /p^h/, /t^h/, /k^h/. Their results revealed that there is similarity in aspiration and VOT between Chinese plosives and English ones. Therefore, we decided to use some other sounds in our experiment. Although plosives are similar in the two languages, there are many other differences between the English and Chinese phonetic systems. One major difference is the existence of many retroflex but few post-alveolar fricatives in Chinese. Some Chinese retroflex fricatives and affricates are similar to their English counterparts, but they are different in terms of place of articulation. Most of the corresponding sounds in English are alveolar or post-alveolar. For our study, four pairs of fricatives and affricates were selected from the two languages. Sounds in each pair differ in place of articulation, but are similar in terms of manner of articulation. Pairs of vowels were also selected.

After comparing all English and Chinese sounds, we finally found four pairs of consonants and six pairs of vowels that could serve our purposes. The four pairs of consonants are /ʃ/, /ʂ/; /ʒ/, /r/; /tʃ/, /t$ʂ^h$/; and /dʒ/, /tʂ/. The former one in each pair is an English sound. Except for the pair /ʒ/-/r/,

sounds in each of the other three pairs are similar but not identical: the three English ones are post-alveolar and the three Chinese ones are retroflex. There is a common mistake made by Chinese English learners' production of [ʒ]. Speakers of Northern Chinese dialects tend to produce a [r] for [ʒ] when there is an l or r in the word. For example, /ʒ/ in word *usual* is often produced as [r]. This phenomenon was mentioned by Ho (2003) as a common mistake produced by people from the Northern China, particularly among Beijing students. Therefore, for the /ʒ/-/r/ pair, tests for contextual effect need to be conducted for Chinese /r/, substituted /r/, and English /r/ (substituted /r/ not included).

The experiment also involves testing vowels. We only tested monophthongs across the two languages. Tested Chinese vowels are /a/, /o/, /i/, /ɛ/ and /u/; and tested English vowels are START, THOUGHT, LOT, GOOSE, FLEECE, TRAP, and DRESS. Mandarin Chinese has six monophthongs, but only five monophthongs were chosen from Chinese because of their similarity to a number of English monophthongs. The Chinese monophthong we excluded from the analysis is /y/, which is located near FLEECE with high F1 and F2 but is a rounded vowel. Some of the Chinese testing examples are /a/ as in *shafa*, /o/ as in *bozi*, /i/ as *yifu*, /ɛ/ as in *yeli* and /u/ as in *shufu*. If the result reveals that any of the vowels shifts toward the other language counterpart, it means that they shift across priming conditions. Speakers of Chinese for whom English is a second language typically find it difficult to distinguish between LOT and THOUGHT (Chang, 2001). TRAP was also tested, because we were keen to see if the participants could separate it from DRESS. The two diagrams below are the English and Chinese vowel charts. Note that Chinese has a high front rounded vowel [y], which is not included in our experiment because there is no corresponding vowel in English, is shown in the diagram. The first diagram is a phonetic vowel space of New Zealand English. The second one is a phonetic vowel space of standard Mandarin, which is based on Beijing standard Chinese.

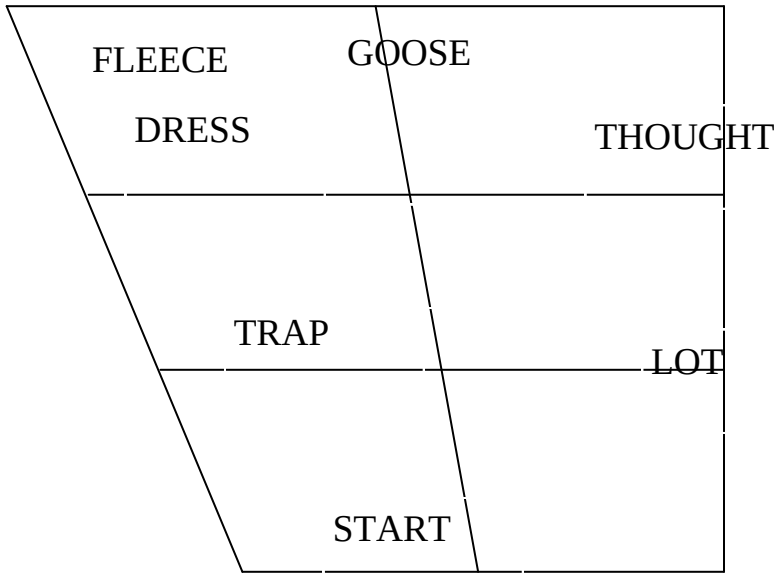

Figure 3. 2: A vowel chart of New Zealand English (Hay, Maclagan & Gordon, 2008)

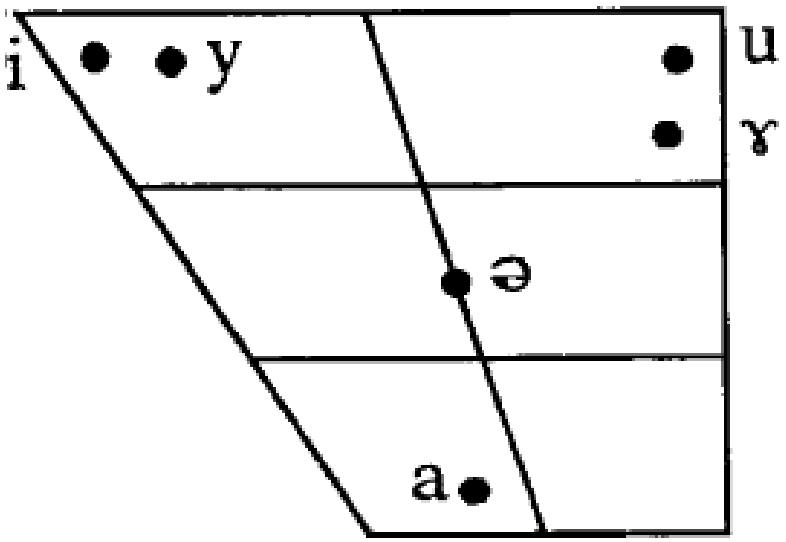

Figure 3. 3: A vowel chart of standard Mandarin Chinese (Lee & Zee, 2003)

3.4.2 Reading passages

Two passages were composed for the participants to read. Each passage contains a list of words that contain all the target sounds, including five consonants, and five-seven target vowels. There are seven-eight tokens for each consonant and four-six for each vowel. Sounds in each comparison across the two passages are matched in syllable positions. Neither the English nor Chinese fricatives have specific positions. English voiced fricatives and Chinese voiced approximants are both syllable-initial in the final syllable. As for Chinese and English affricates, these are located mostly at word-initial and word-final positions, but some in word-medial position. All the vowels tested occur in stressed syllable after plosives, fricatives and affricates. None of the vowels appears before lateral consonants, so as to avoid prelateral effects. As has been mentioned, the first passage participants read was in the same language as that of the questionnaire, and the second passage was in the other language. The passages were checked by native speakers to ensure their coherence.

3.4.3 Data measurement

All the reading passages were transcribed and time-aligned by hand by using Transcriber, automatically segmented by using the HTK toolkit, and then target segments were hand-corrected in Praat. Vowel duration was taken from the start of showing formants to the end of the formants or to the start of the following consonant if it has one. Consonants were measured by taking the entire duration of frication, i.e. the noise part.

Vowels

After transcribing all the sound files and aligning a textgrid for each sound file, praat scripts were written and run to automatically extract measurements[1]. Vowels were measured by taking the first and second formant at the centre of the vowel.

Consonants

The measurement of fricatives and affricates can be done in various ways. Normally, Centre of Gravity (COG), which is the mean value of signal weight and can show the energy of a sound, is useful for measuring fricatives, because it can differentiate certain fricatives (Gordon, Barthmaier & Sands, 2002). Gordon, Barthmaier & Sands' (2002) cross-language study of fricatives of seven non-English languages showed that centre of gravity could distinguish most of their fricatives. Li et al. (2007) present a new way of measurement. They use a two dimensional graph to separate dental alveolar /s/, retroflex /ʃ/ and alveolopalatal /ɕ/ by taking an 'ampRatio', which is the difference of amplitude between the highest peak and F2, and the centroid frequency, which is the mean frequency of a spectrum. Halle and Stevens (1997) point out that the difference between a alveolopalatal fricative and a retroflex fricative is a spectral peak of frication spectrum at around 1500Hz for the alveolopalatal fricative. In the present study, we do not take the centre of gravity as a parameter. According to observation of the recordings, frication produced by Chinese speakers in the current study all has a very low band of energy at the bottom of its noise spectrum. The presence of these low frequencies would affect the overall average value of the

[1] Thanks to Jennifer Hay, who is the senior supervisor for the thesis, for writing the script for me.

centre of gravity. The graph below is a comparison of /ʃ/ in the word ‘shelf’ produced by a female Chinese speaker (speaker 9 in English context) from the experiment and the word ‘share’ produced by a female native English speaker from Southland. The centre of gravity on the left is only 950.2426 Hz, whereas the centre of gravity of /ʃ/ produced by a native speaker is 2137.3474 Hz. The frequencies from the two words look otherwise similar but the presence of the low frequency activity makes the COG non-comparable.

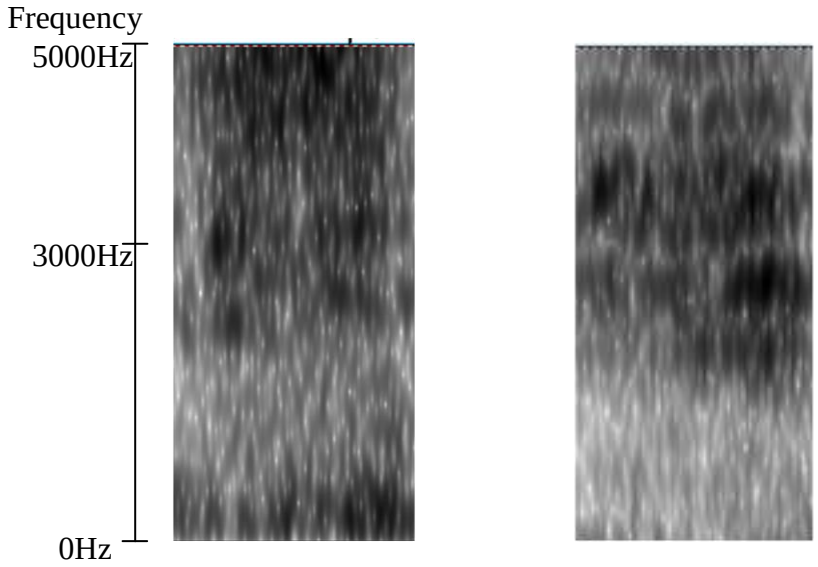

Figure 3. 4: /ʃ/ as in ‘shelf’ (left) and in ‘share’

With low energy in the entire frication, the value of centre of gravity would not accurately represent the characteristics of frication produced by the Chinese speakers. Despite the fact that the COG at higher than 1500Hz would be accurate, we tend to use another type of measurement. The low energy might in fact be seen as a distinction in voicing. Our speakers tend to create a voice bar during the fricatives. Such a phenomenon is also found by Holton (2001) in Tanacross Athapaskan voiced fricatives and semi-voiced fricatives. In Chinese, a semi-voiced /ʃ/ is indeed found from some dialectal regions (Ho, 1996), which might suggest that our Chinese participants can produce semi-voiced fricatives, too. The two graphs below show the difference of the two spectrums.

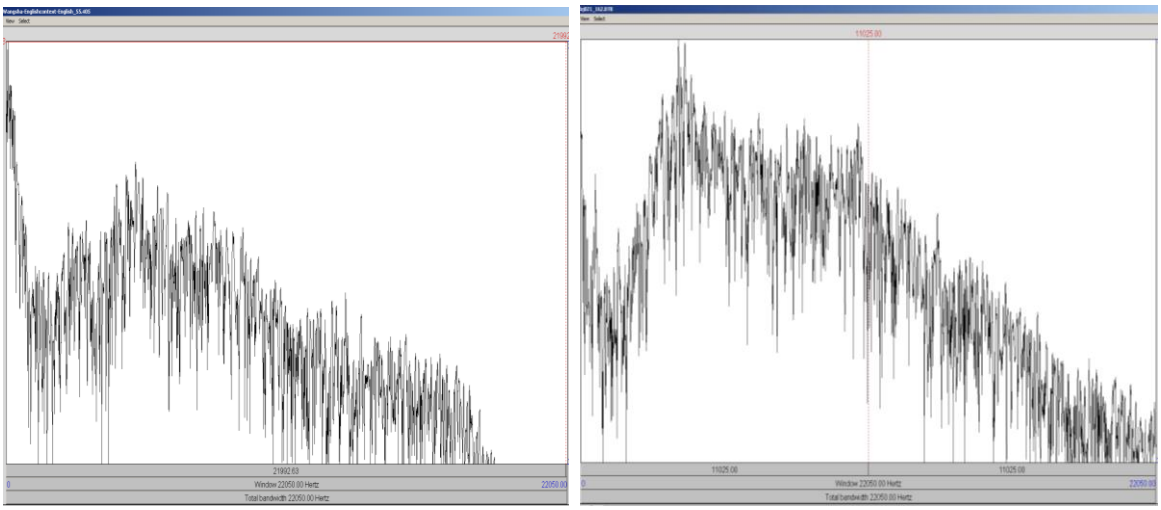

Figure 3. 5: Spectrum of /ʃ/ as in ‘shelf’ produced by the Chinese female speaker (left) and spectrum of /ʃ/ as in ‘share’ produced by the English female speaker

The comparison shows an interesting tendency: while producing frication, the Chinese speaker’s production has higher amplitude at zero frequency than the peak at the frequency higher than 1000Hz. Such a fact could be caused by different voice qualities. In contrast, in the English speaker’s production, the amplitude at the frequency higher than 1000Hz is actually the real peak, so it is higher than the amplitude at zero frequency. More interestingly, in another speaker’s production, the second peak is higher amplitude than the first peak at the frequency higher than 1000Hz. The spectrum is given below after LPC smoothing.

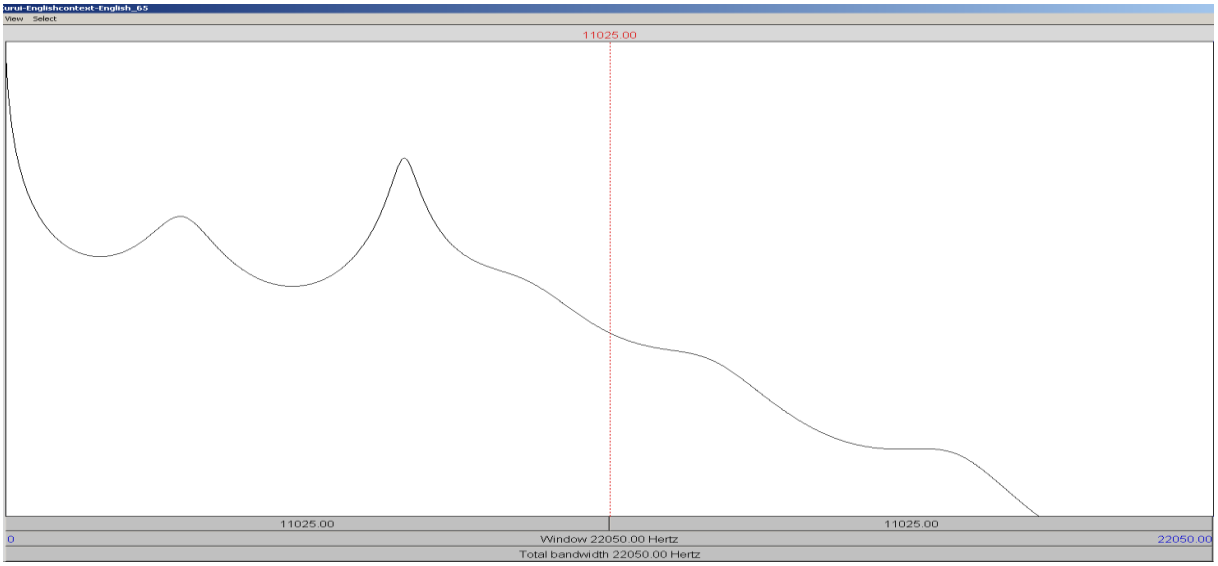

Figure 3. 6: LPC smoothing spectrum of /ʃ/ as in ‘shelf’ produced by a male Chinese speaker

It seems that the second peak is the peak of interest for establishing place of articulation. The speaker who produces frication with two peaks consistently produces frication in the same way throughout the two sections in our experiment.

The high amplitude at zero frequency raises the question of whether the fricatives are more voiced in the English or Chinese context. Chinese retroflex fricatives are voiceless sounds, but spectrums suggest that there might be different voice qualities when they are produced by some of our speakers.

As a result, in the present study, we attempted to take two measurements from speakers’ fricatives. Measurements are not taken from the entire frication, but at the centre fifty percent of the frication from each fricative and affricate, so that the data are less likely affected by surrounding segments. Then we created a ‘cepstral smoothed’ spectrum over this portion. We measured amplitude difference (ampdiff or Amp Diff) between the amplitude at zero frequency and the highest peak at the frequency higher than 1000Hz, following the observation that this was often positive in the Chinese speech, but negative in the English speech. The second measurement we took was the frequency at the highest peak above 1000Hz. This was an attempt to gauge the relative place of articulation, while eliminating the lower frequencies which are problematic for reasons outlined above. For affricates, we used the same method of measurement.

In following the procedures of the experiment, each participant produced recorded data in both languages in the environment of both contexts. All the tested segments were produced in the immediate language environment, thus we could analyze contextual effects of short-term exposure on both Chinese and English segments, and discover the effects between L1 and L2 in the speech of our participants.

Chapter 4

Results of Vowels

One important aim of the present study is to see if speakers' L1 changes in L2 context as a result of L2 proficiency. Following that, we would also like to see if speakers' L2 changes in L1 context as well. The experiment first needs to compare English vowels and consonants with their Chinese counterparts. Then, data analyses will examine if any Chinese segments have moved toward their English counterparts, and vice-versa. The hypotheses are as follows.

Hypotheses:

The effect of the English context should result in changes of pronunciation in Chinese and vice versa. Changes can be found within pairs of corresponding sounds in the two languages. If a

vowel in L1 is affected by L2, results should indicate that values of the L1 vowel would move closer to its L2 counterpart either in F1 or F2 in the L2 context. If short-term priming effects were correlated with L2 proficiency, we would see speakers from different proficiency groups perform differently as a result of priming.

4.1: Contextual Effects

Tested vowels in this experiment are /a/, /i/, /e/, /u/ and /o/ in Chinese, and /a/, /i/, /e/, /æ/, /u/, /o/ and /ɔ/ in New Zealand English. Each tested Chinese vowel can find its approximate counterpart in English. Two English vowels do not exist in Chinese: /æ/ and /ɔ/. Table 4.1 gives each tested vowel and their token numbers.

English vowel	number of tokens	Chinese vowel	number of tokens
START/ a/	6	/a/	5
LOT/ɔ/	4	/e/	5
TRAP/æ/	5	/i/	6
DRESS/e/	5	/o/	5
FLEECE/i/	6	/u/	5
THOUGHT/o/	4		
GOOSE/u/	6		

Table 4. 1: Tested vowels in each language and their token numbers

One aim of the experiment was to find out possible effects of the English context on Chinese sound production. Therefore, the first thing needs to do is to compare the Chinese tokens in the two contexts. After that, if any (near) significant can be found, we need then to identify the differences between the Chinese and English vowels produced by these speakers, which can help provide an explanation for any contextual differences found, e.g. transfer effects. Wilcoxon-tests for male speakers are carried out by using R. After comparing each Chinese vowel across different contexts, results show that apart from F2 of /u/, which gives a near significant result $p = 0.09451$, the other four vowels are not produced differently across the two contexts. When comparing Chinese vowels with English vowels, the results show that most of the Chinese

vowels are produced significantly differently from their English counterparts. The results are given in Table 4.2

Chinese	English	Formant	p-value
/i/	FLEECE	F1	<0.0001
/i/	FLEECE	F2	0.7857
/e/	DRESS	F1	<0.0001
/e/	DRESS	F2	0.05598
/e/	TRAP	F1	<0.0001
/e/	TRAP	F2	<0.0001
/u/	GOOSE	F1	<0.0001
/u/	GOOSE	F2	<0.0001
/a/	START	F1	<0.0001
/a/	START	F2	0.000157
/o/	LOT	F1	<0.0001
/o/	LOT	F2	0.3073
/o/	THOUGHT	F1	<0.0001
/o/	THOUGHT	F2	0.03283
/o/	LOT + THOUGHT	F1	<0.0001
/o/	LOT + THOUGHT	F2	0.0631

Table 4. 2**:** Differences between Chinese vowels and their English counterparts

Results in the table above provide reasonable guidelines for testing the contextual effect on L1: if Chinese vowels were identical to the English ones, it would be meaningless to test the effect of an English context on Chinese vowels, and vice versa.
Regression models were then used to test each pair of vowels listed in Table 4.2 that were produced significantly differently across the two languages. I first present the Chinese results.

4.1.1: Chinese vowel production

If interference between the two language systems occurs, Chinese vowel production should also be affected by language context. Results show that Chinese /u/ is produced almost differently across the two contexts. Regression models were calculated for both F1 and F2 of each Chinese

vowel, and proficiency is included in the models. Table 4.3 gives p-values for context coefficient when context and proficiency are tested as separated fixed effects; whereas Table 4.4 shows results of models when testing the interaction.

p-value	/u/	/i/	/e/	/a/	/o/
F1	0.55	0.3981	0.3258	0.1358	0.8497
F2	0.0374	0.1744	0.9357	0.3577	0.3119

Table 4. 3: Context coefficient of regression models of testing context and proficiency as separated effects

p-value	/u/	/i/	/e/	/a/	/o/
F1	0.5144	0.2526	0.1186	0.8364	0.8041
F2	0.1697	0.7627	0.6836	0.2564	0.4942

Table 4. 4: Context coefficient of regression models when testing the interaction

Results in Table 4.3 reveal that for Chinese /u/, its F2 is influenced by the different language contexts, and is affected by L2 proficiency. This result is consistent with the result in the comparison of /u/ across the two contexts. The prediction is not borne out by the other Chinese vowels tested in the experiment.

F2 of GOOSE and Chinese /u/

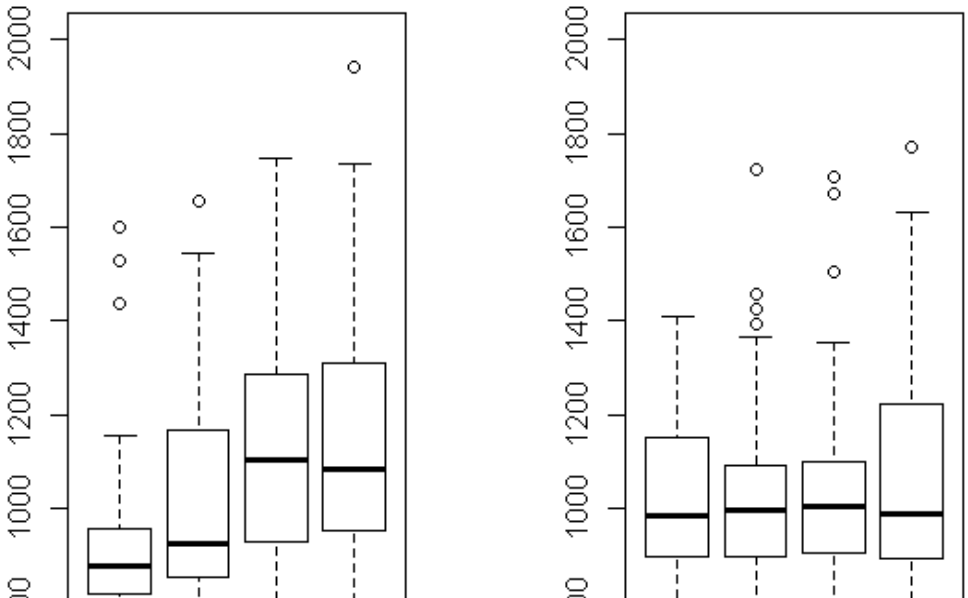

Figure 4. 1: Comparison of English GOOSE and Chinese /u/ in different contexts across low and high proficiency speakers (e.g. hec = high proficiency speaker, English context and Chinese production)

As Figure 4.1 shows, English GOOSE has an overall higher F2 than Chinese /u/. More importantly, high proficiency speakers make a great distinction between GOOSE and Chinese /u/. Table 4.5 shows that the English context can increase F2 of Chinese /u/ (by 60.81Hz) for both proficiencies, which means producing a more English-like vowel.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	941.83	942.22	822.058	1056.2	0.0001	0
Proficiency=low	58.26	57.94	-75.979	185.8	0.3634	0.4543
Context=English	60.81	60.64	1.759	120.1	0.0448	0.0374

Table 4. 5: Results of a regression model for predicting F2 of Chinese /u/

Chinese /a/

Although Chinese /a/ was not found to be significantly affected by the different contexts, the way speakers moved /a/ in the English context is worth mentioning.

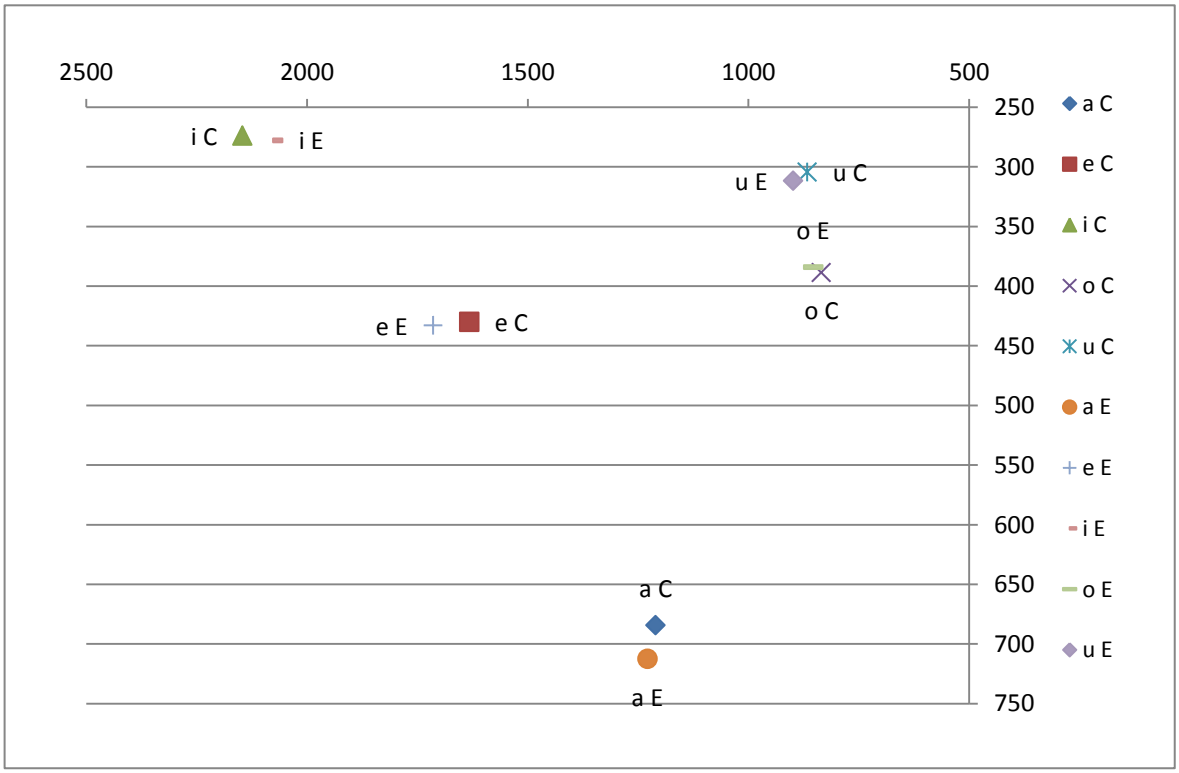

Figure 4. 2: Chinese vowels produced by male high proficiency speakers across different contexts (C = Chinese context, E = English context)

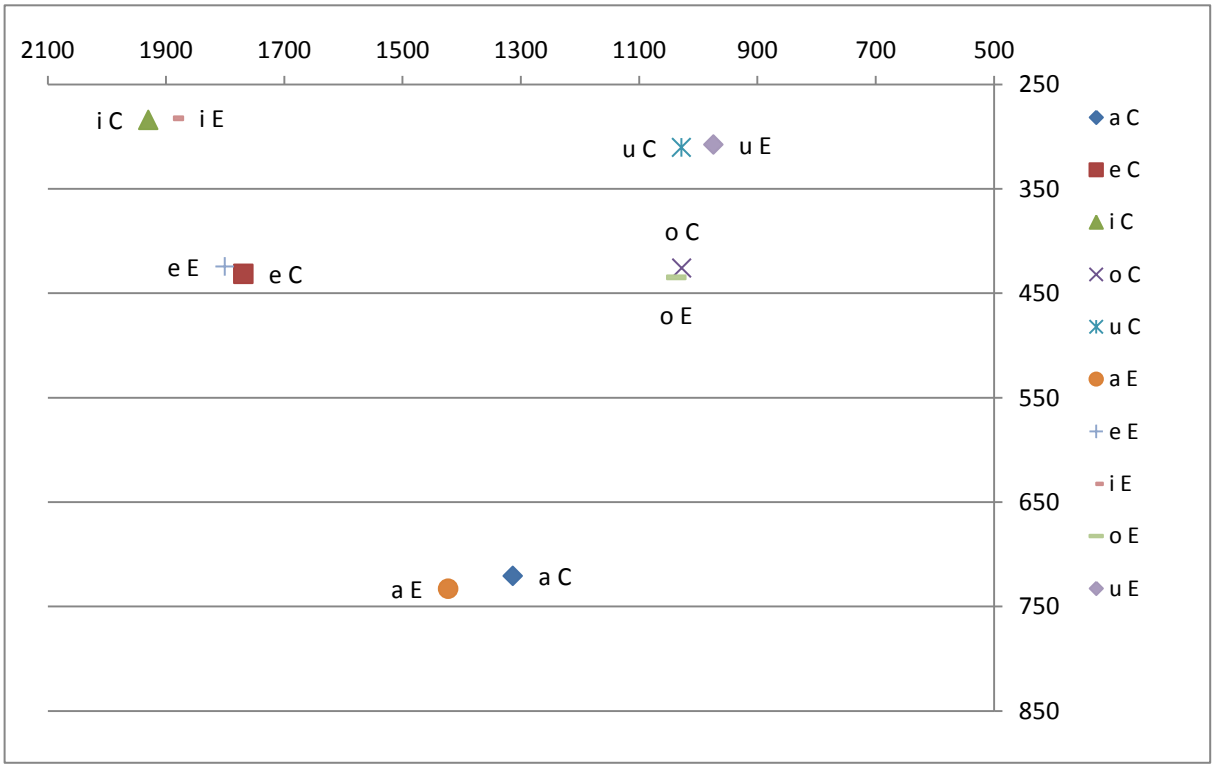

Figure 4. 3: Chinese vowels produced by male low proficiency speakers across different contexts

Figure 4.2 and 4.3 are Chinese vowels produced by male speakers. High proficiency speakers separate /a/ in the English context away from that in the Chinese context by increasing its F1. Low proficiency speakers increase the F2 of /a/ in the English context as well. The table below gives mean values of /a/ in both contexts produced by speakers of both proficiencies.

F1(Hz)	F2(Hz)	context	proficiency
684.3333	1210.233	Chinese	high
712.2759	1228.379	English	high
720.7391	1313.087	Chinese	low
732.7778	1422.444	English	low

Table 4. 6: Mean F1 and F2 of Chinese /a/ produced by male speakers in different contexts

It will be interesting to see if such patterns can be transferred to their English production.

4.1.2: English vowel production

For each English vowel, firstly, both F1 and F2 were compared across the two contexts, however, no significant differences were found. After that, both F1 and F2 were tested in a mixed-effect regression model with 'speaker' and 'word' as random effects. Main tested fixed effect is context; however proficiency is also included in the model as another possible factor. F1 and F2 for each vowel were tested in two different models: one included context and proficiency as separated fixed effects, whereas the other one tested the interaction of context and proficiency. P-values for the 'context' coefficients in models are shown in Table 4.7 and 4.8. The first table includes results from models which have separated fixed effects; whereas the second table is from models which test the interaction. Amongst all of the test results, three vowels show a (near-) significant effect of context on English vowels: English LOT, THOUGHT, and START.

p-value	FLEECE	DRESS	TRAP	GOOSE	START	THOUGHT	LOT
F1	0.6132	0.5786	0.2767	0.6223	0.8197	0.6912	0.079
F2		0.5786	0.1264	0.8101	0.1486		0.5377

Table 4. 7: P-value of regression models for predicting F1 and F2 of English vowels when considering 'context' and 'proficiency' as separated factors (the empty space represents the absence of significant difference in F2 between an English vowel and its Chinese counterpart)

p-value	FLEECE	DRESS	TRAP	GOOSE	START	THOUGHT	LOT
F1	0.5794	0.2108	0.8195	0.1494	0.5636	0.0566	0.3443
F2		0.3701	0.8339	0.9148	0.0607		0.6121

Table 4. 8: P-value of regression models for predicting F1 and F2 of English vowels when testing the interaction of 'context' and 'proficiency'

LOT and THOUGHT

The interaction model uses the interaction of context and proficiency as a fixed effect. It is based on the assumption that contextual effect might be different across different proficiency groups. For example, the result of F1 for THOUGHT from the interaction model is near significant (Table, 4.9), which suggests that low proficiency speakers tend to decrease (-55.41) F1 of THOUGHT in the English context. In contrast, in Table 4.10, the near significant result is only

from context, which is not proficiency specified, suggesting that both high and low proficiency groups have the tendency to decrease (-25.67) F1 of LOT in the English context.
Tables 4.9 and Table 4.10 show near significant results for testing THOUGHT and LOT.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	486.96	487.13	427.3	550.37	0.0001	0
Proficiency=low	60.53	60.56	-12.04	138.185	0.1138	0.2459
Context=English	18.23	18.41	-22.62	59.525	0.3862	0.3373
Proficiency=low: Context=English	-55.41	-56.24	-120.42	5.277	0.0766	0.0566

Table 4. 9: Results of a regression model for predicting F1 of English THOUGHT

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	501.03	500.9	449.06	556.27	0.0001	0
Proficiency=low	31.93	32.1	-30.68	96.629	0.3046	0.4463
Context=English	-25.67	-25.7	-55.99	5.038	0.1012	0.079

Table 4. 10: Results of a regression model for predicting F1 of English LOT

Note that R picks factors alphabetically. Negative value for the English context indicates an increase in F1 of THOUGHT in the Chinese context. Low proficiency speakers in this study have a tendency to produce THOUGHT with high F1 in the Chinese context (p= 0.079).
THOUGHT is not near a vowel in Chinese. /o/ is the closest vowel. The comparison can be seen between Figure 4.4 and Figure 4.5.

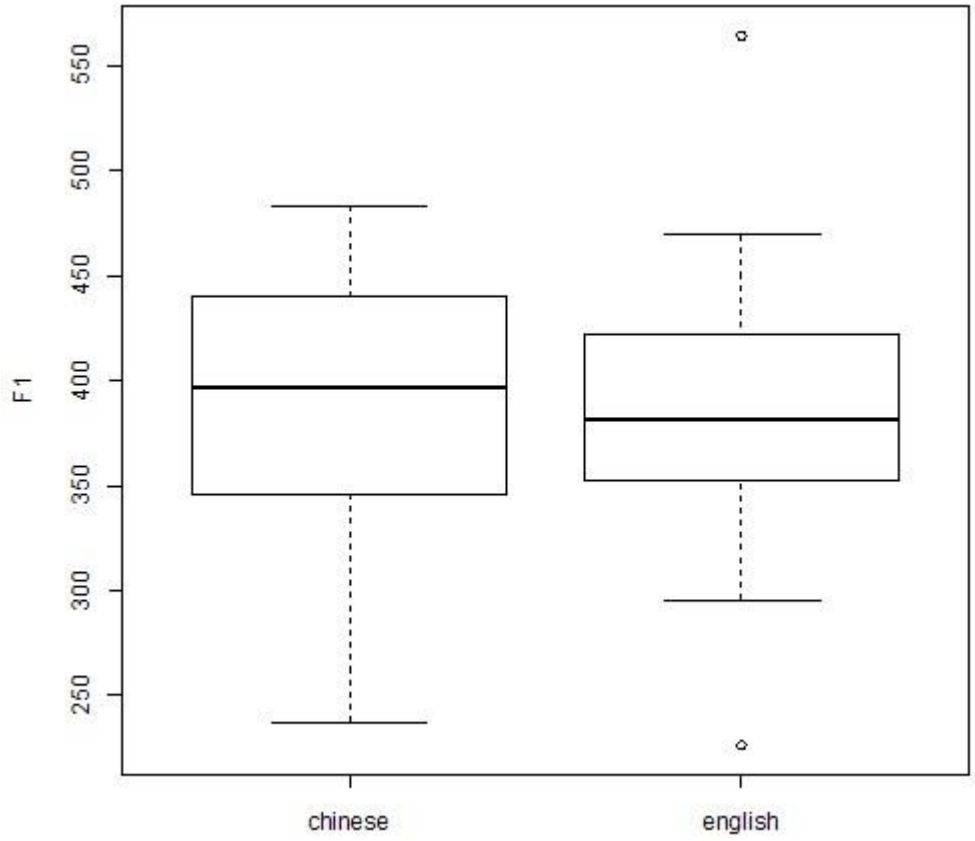

Figure 4. 4: F1 of Chinese /o/ produced by high proficiency speakers in the Chinese (left) and the English (right) context

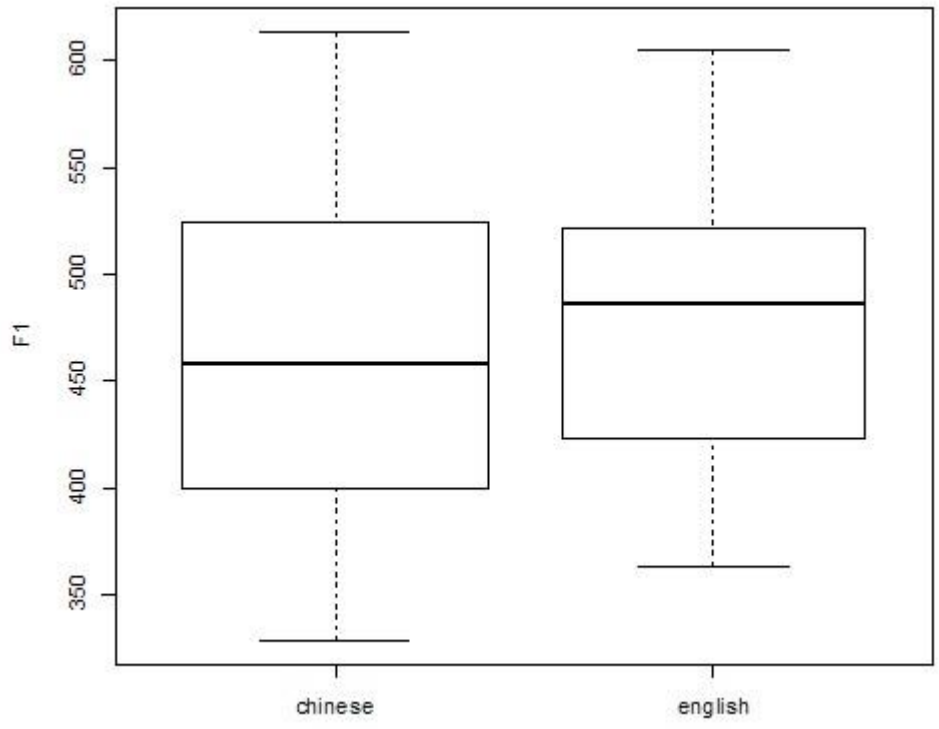

Figure 4. 5: F1 of English THOUGHT produced by high proficiency speakers in the Chinese (left) and the English (right) context.

The main purpose of the two graphs is to compare English and Chinese, rather than to show or discuss contextual effects in Chinese. Mean F1 of English THOUGHT is about 450~500Hz. In contrast, mean F1 of Chinese /o/ is about 350~400Hz. Figure 4.6 shows a comparison between English THOUGHT and Chinese /o/ across contexts by both low and high proficiency speakers.

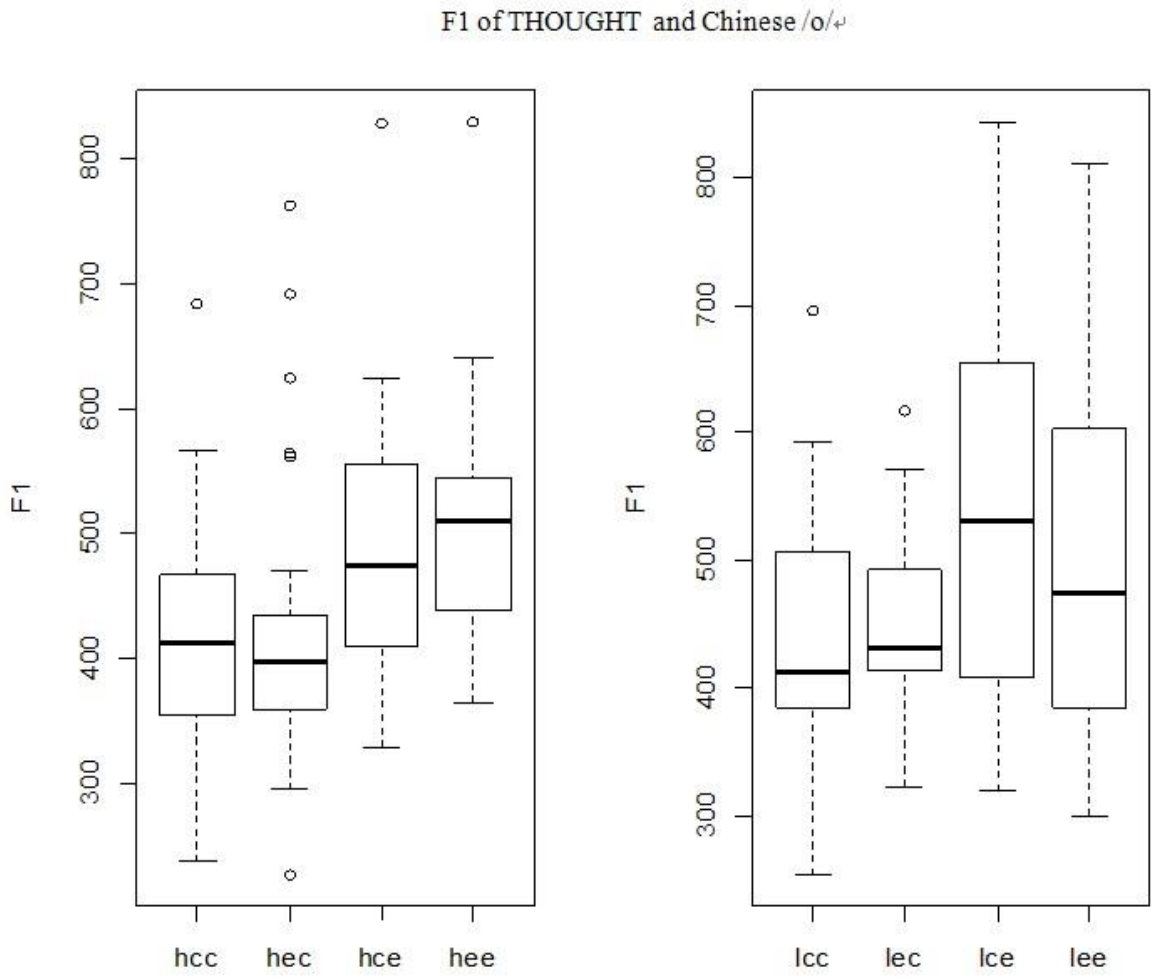

Figure 4. 6: Comparison of English THOUGHT and Chinese /o/ produced in different contexts across low and high proficiency speakers (e.g. hec = high proficiency speaker, English context and Chinese production)

Figure 4.6 shows that both low and high proficiency speakers produce lower THOUGHT (higher F1) than Chinese /o/. But that there is only a marked contrast across contexts in the English THOUGHT for the low proficiency speakers. This is consistent with the statistical results. Unexpectedly, they are producing a more English-like vowel when primed with the Chinese context.

Table 4.8 gives a near significant result for LOT. It shows that the English context decreases (Chinese context increases) F1 of LOT for all the speakers. The results in Figure 4.7 and 4.8 show that F1 of Chinese /o/ is lower than that of English LOT.

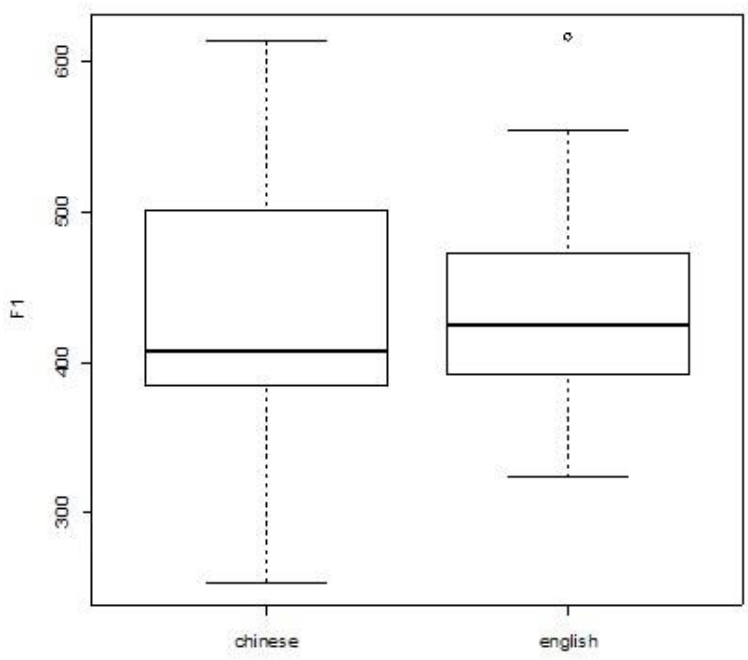

Figure 4. 7: F1 of Chinese /o/ produced by low proficiency speakers in Chinese (left) and English (right) context.

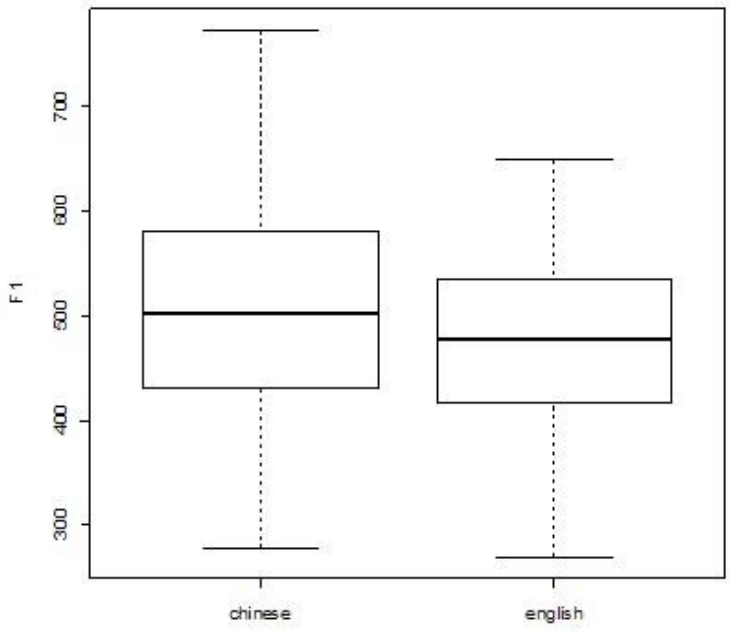

Figure 4. 8: F1 of English LOT produced by low proficiency speakers in Chinese (left) and English (right) context.

Again, the results for the Chinese context are only included in the graph for comparison. As has been shown above, mean F1 of Chinese /o/ is about 400~450Hz, but mean F1 of English LOT is about 450~500Hz.

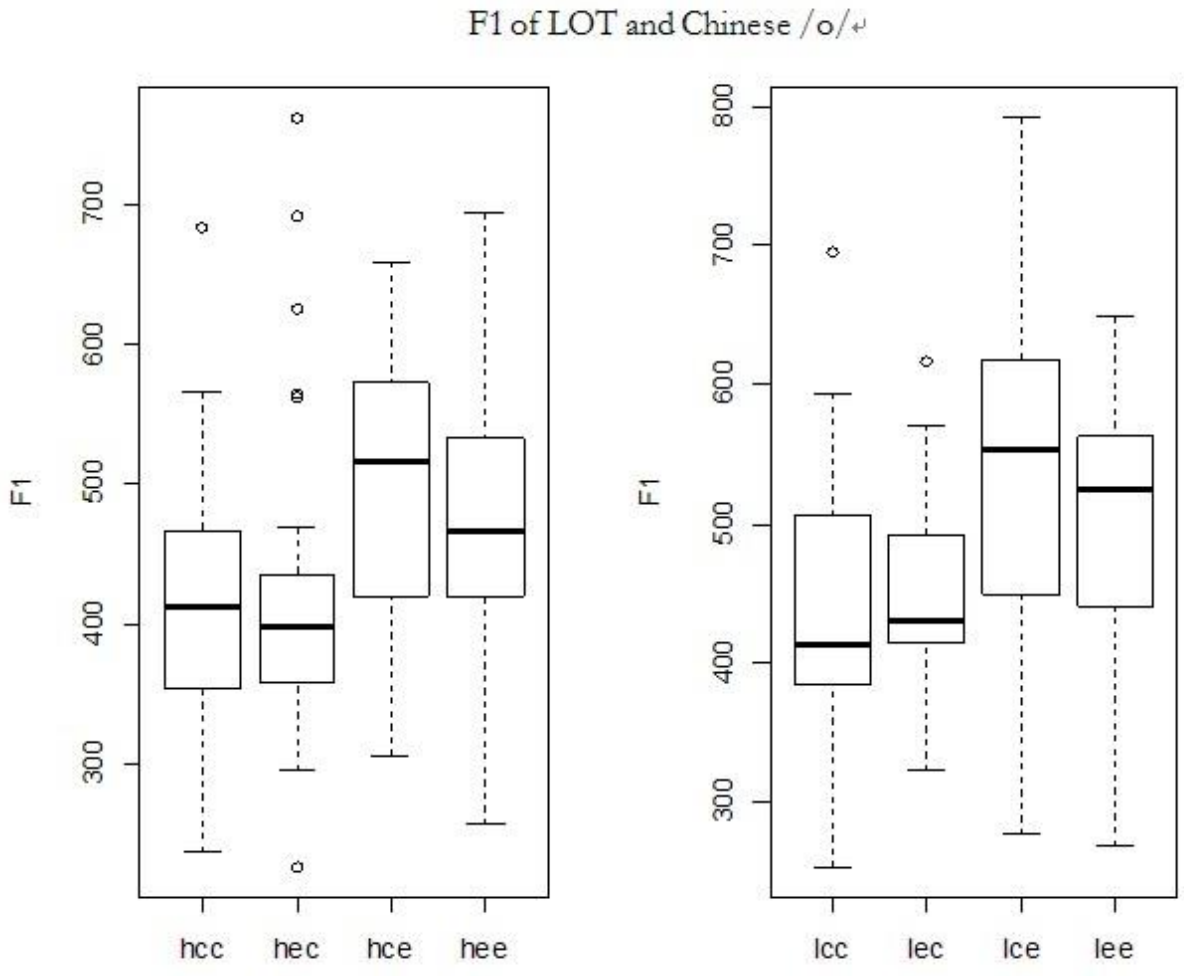

Figure 4. 9: Comparison of English LOT and Chinese /o/ in different contexts across low and high proficiency speakers (e.g. hec = high proficiency speaker, English context and Chinese production)

Both low and high proficiency speakers produce LOT with higher F1 than Chinese /o/. Overall results of /o/ and LOT indicated that for both proficiency speakers, the English context will decrease F1 for both of them. In other words, because F1 of Chinese /o/ is lower than LOT, the decrease leads LOT move to the direction of its Chinese counterpart in the English context.

START

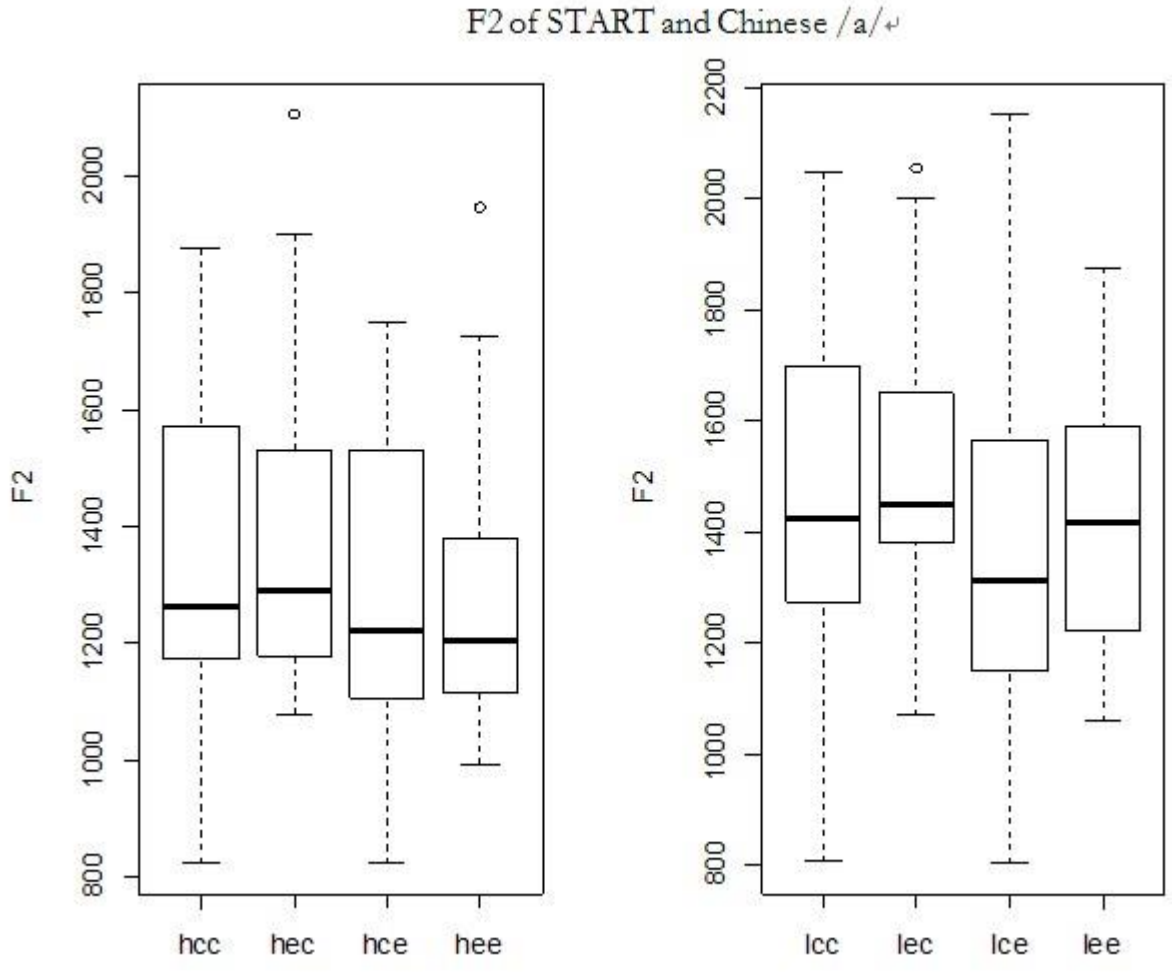

Figure 4. 10: Comparison of English START and Chinese /a/ in different contexts across low and high proficiency speakers (e.g. hec = high proficiency speaker, English context and Chinese production)

Figure 4.10 illustrates a comparison between English START and Chinese /a/ in F2. It reveals that F2 of Chinese /a/ is higher than that of START. The near significant result for START is shown below:

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	1290.219	1289.992	1167.049	1415.42	0.0001	0
Proficiency=low	51.132	49.402	-107.537	198.27	0.528	0.6607
Context=English	-3.914	-5.135	-73.689	64.76	0.891	0.9045
Proficiency=low: Context=English	94.516	100.527	-5.562	205.8	0.0664	0.0607

Table 4. 11: Results of a regression model for predicting F2 of English START

The result shows that when low proficiency speakers speak English, the English context increases F2 of their English START, making it closer to the Chinese vowel.

In summary, when speaking English, the immediately preceding context does not significantly affect any of the vowels, but there are near-significant effects in three cases. Counter-intuitively, in all three cases, the effect of the English language prime is to produce more Chinese-like vowels. In two of the three cases, the trend is only present for the low proficiency speakers.

Analyses of the three vowels

Both high and low proficiency speakers merged LOT and THOUGHT. Low proficiency speakers' productions of the two have lowered F1in both contexts. The cross context comparison is given in Figure 4.11.

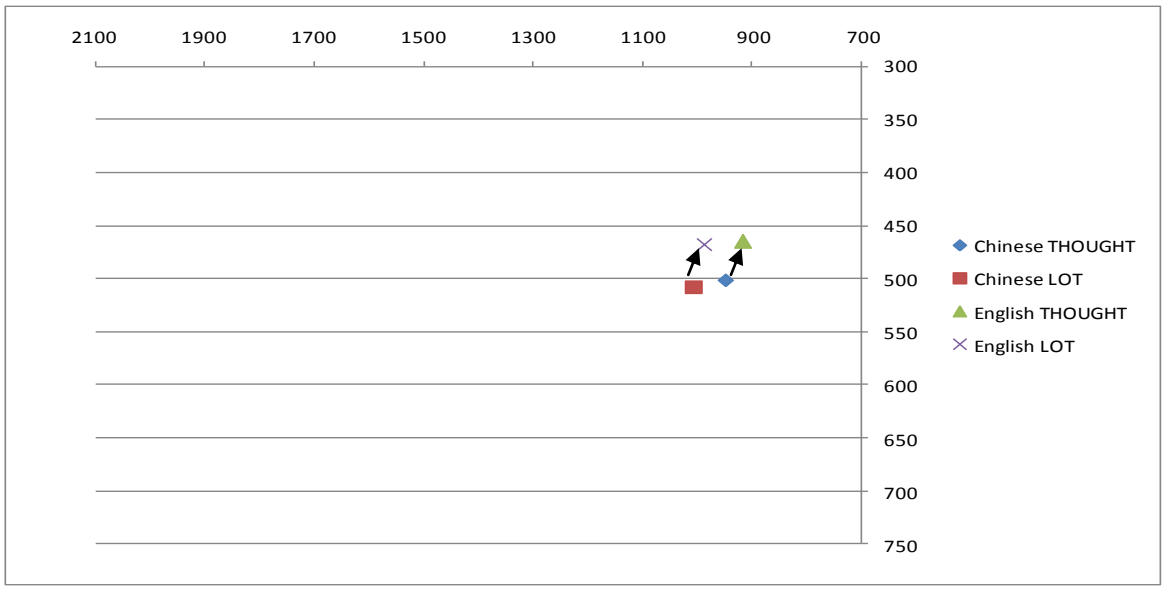

Figure 4. 11: Comparison of THOUGHT and LOT in the English context (top two) and the Chinese context for low proficiency speakers

The figure shows that low proficiency speakers produce LOT and THOUGHT with lower F1 in the English context than in the Chinese context. It might be plausible that low proficiency speakers actually can have some conscious awareness of the difference between English LOT and THOUGHT vowels; that is, of the fact that they should be distinct. A reasonable interpretation could be that when they produce the two vowels in the English context, the environment raises the awareness that the two should be produced differently; however, since THOUGHT vowel has no counterparts in Chinese, they move both vowels in the direction of Chinese /o/. Figure 4.6 and 4.9 reveal that low proficiency speakers' F1 of both THOUGHT and LOT decreases in the English context. Secondly, recall that high proficiency speakers separate /a/ in the English context away from that in the Chinese context by increasing its F1, but low proficiency speakers increase the F2 of /a/ in the English context. When speaking English, low and high proficiency speakers seem to use the same two strategies to separate START from LOT and THOUGHT.

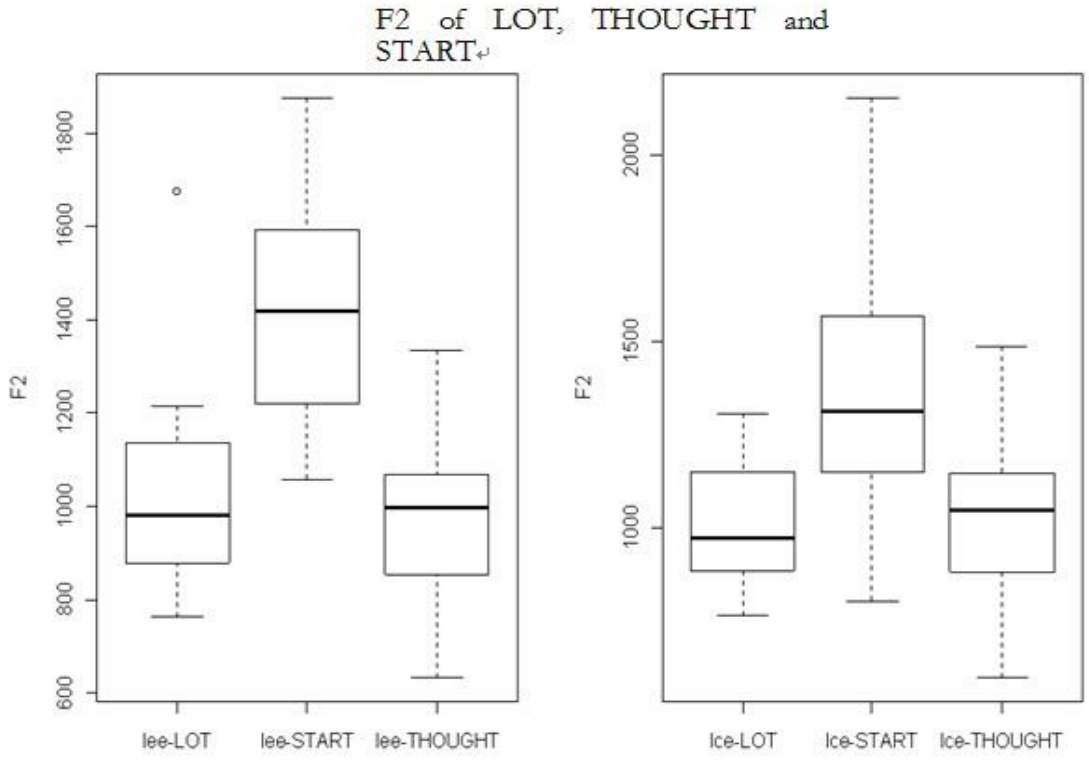

Figure 4. 12: Comparison of F2 in LOT, START and THOUGHT produced by low proficiency speakers across contexts (e.g.lce = low proficiency speaker, Chinese context and English production)

Figure 4.12 indicates that low proficiency speakers use F2 to separate START away from LOT and THOUGHT in the English context. Note that in the English context, the difference of F2 between START and LOT-THOUGHT is bigger than that in Chinese context.

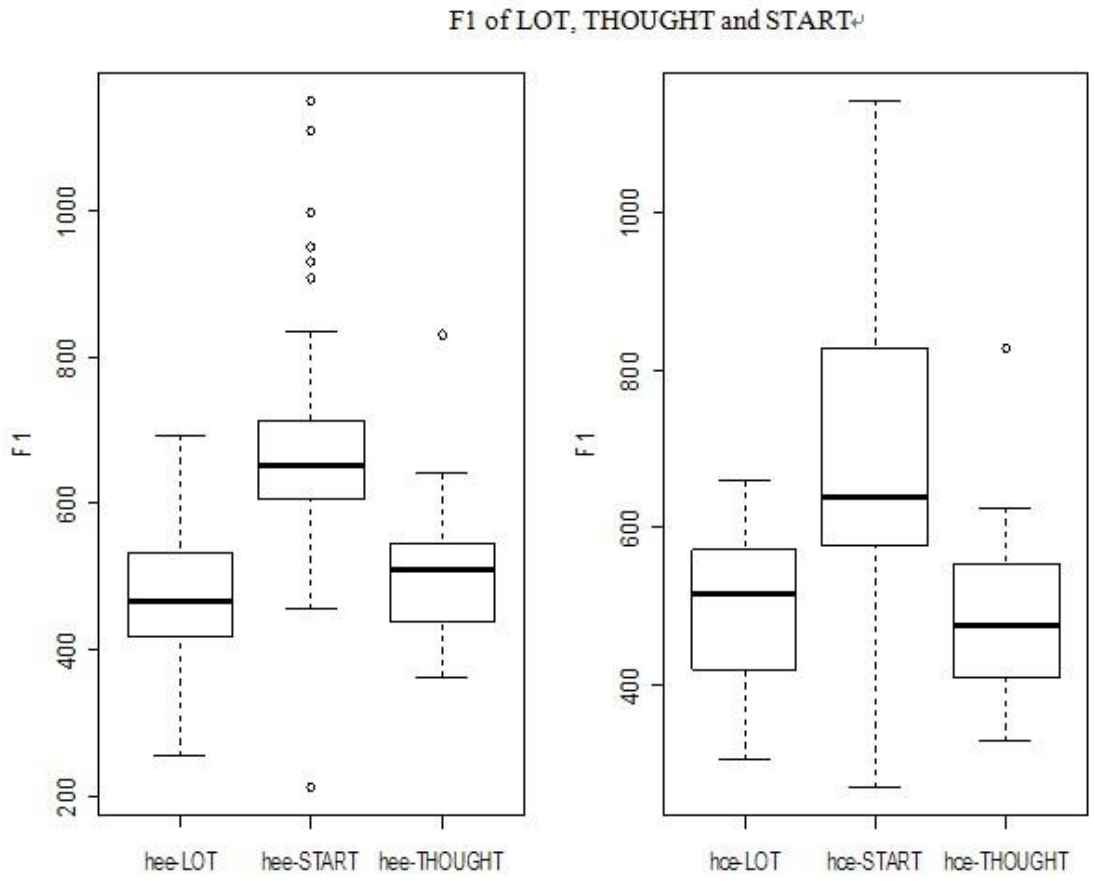

Figure 4. 13: Comparison of F1in LOT, START and THOUGHT produced by high proficiency speakers across contexts (e.g. hce = high proficiency speaker, Chinese context and English production)

Figure 4.13 indicates that high proficiency speakers use F1 to separate START away from LOT and THOUGHT in the English context. Note that in the English context, the difference of F1 between START and LOT-THOUGHT is bigger than that in the Chinese context. Graphs below

show distribution of English vowels in both contexts for high and low proficiency speakers. When comparing Figure 4.14 with 4.15, the latter shows that low proficiency speakers' F1 of START has little difference between contexts, however, its F2 shows a bigger distinction. This indicates that low proficiency speakers' F2 of START is more separated out from F2 of LOT and THOUGHT in the English context than in the Chinese one. On the contrary, Figure 4.14 reveals a bigger difference in F1 than in F2 of START for high proficiency speakers. High proficiency speakers tend to use F1 more to separate START away from LOT and THOUGHT in the English context than in the Chinese context.

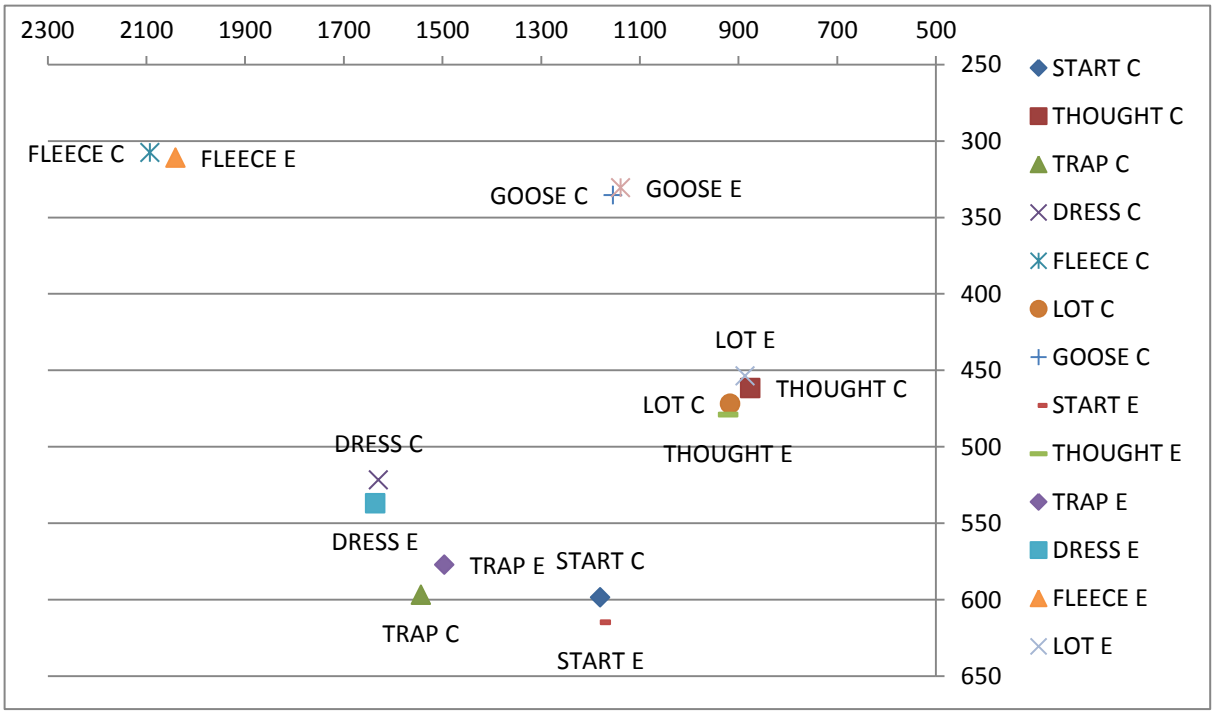

Figure 4. 14: English vowels produced by male high proficiency speakers across different context (C = Chinese context, E = English context)

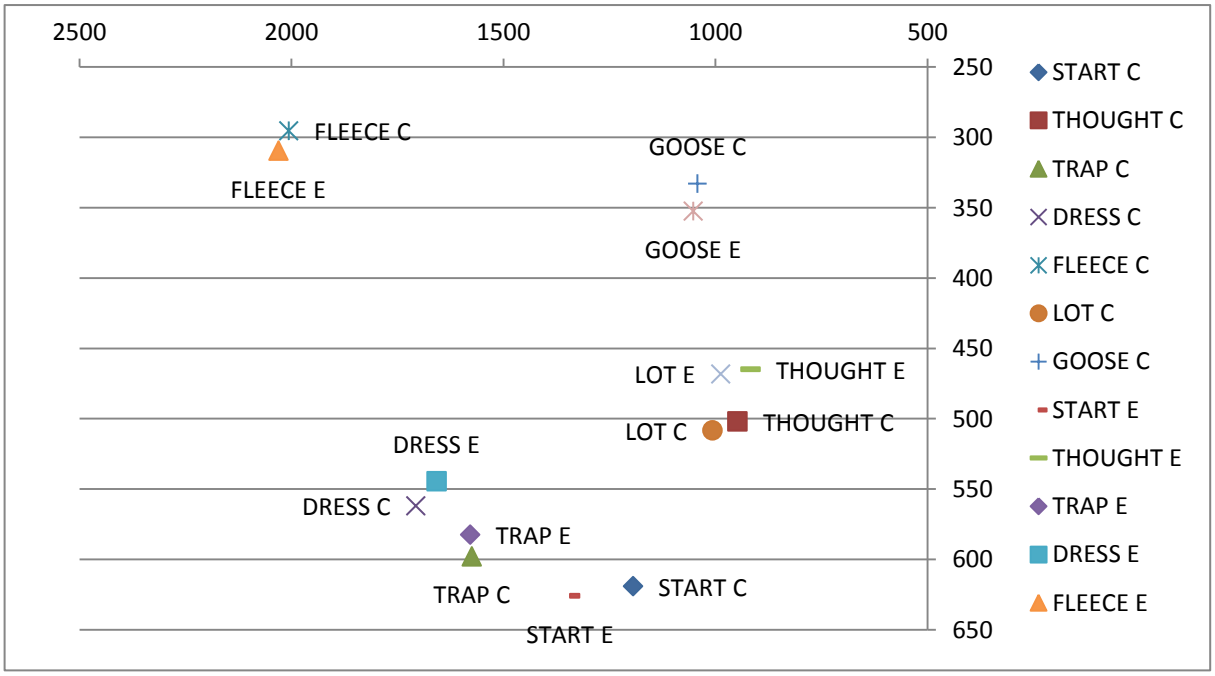

Figure 4. 15: English vowels produced by male low proficiency speakers across different contexts

4.2: Performance of English Vowels

Figure 4.16 and 4.17 demonstrate the distribution of English vowels produced by male low proficiency speakers in the Chinese and the English context. This section contains the participants' production of tested English targets in terms of different characteristics caused by different proficiencies.

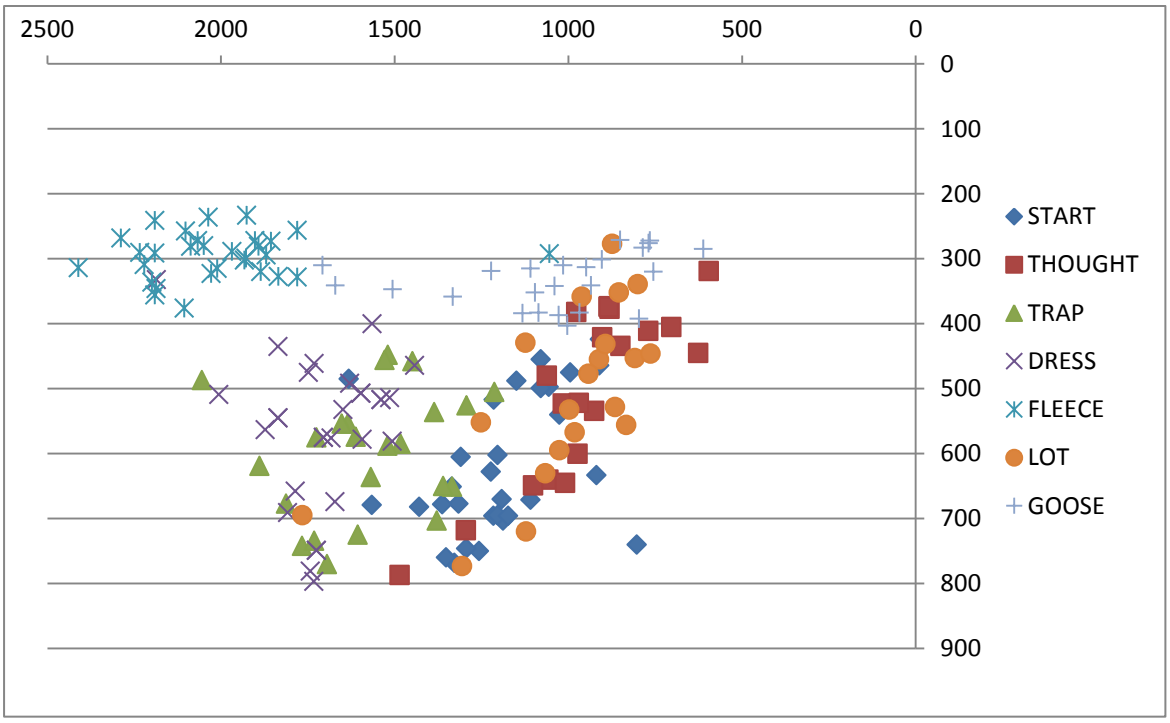

Figure 4. 16: English vowels produced by male low proficiency speakers in the Chinese context

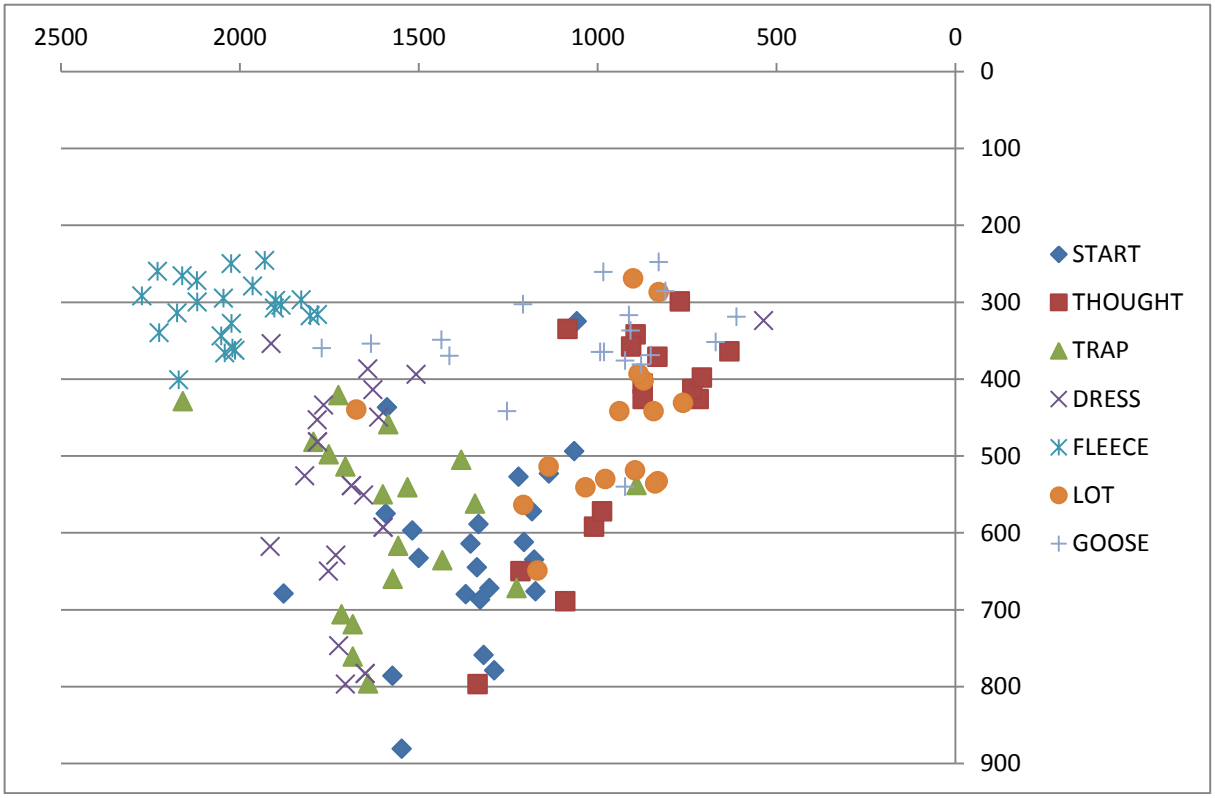

Figure 4. 17: English vowels produced by male low proficiency speakers in the English context

Male low proficiency speakers' F1 of English TRAP vowel and DRESS vowel are almost mixed. LOT vowel and THOUGHT vowel are also not separated. As has been mentioned, TRAP vowel and THOUGHT vowel do not have any approximate counterparts in Chinese. It is thus expected for low proficiency speakers that they hardly produce TRAP and DRESS differently, or LOT and THOUGHT. Non-paired Wilcoxon-tests were done to calculate the differences within each pair by using R, and the results are shown in Table 4.10.

	Chinese context		English context	
	F1	F2	F1	F2
TRAP and DRESS	0.2217	0.0322	0.3	0.05012
LOT and THOUGHT	0.642	0.724	0.4856	0.4623

Table 4. 12: Results of p-value in Wilcoxon-test between TRAP and DRESS, and LOT and THOUGHT across contexts

The only significant results are the difference between TRAP and DRESS in F2 in both contexts, p equals 0.0322 and 0.05012. The other six results are not significant. Male low proficiency speakers' LOT and THOUGHT are completely merged, but TRAP and DRESS are slightly separated in both contexts. Figure 4.18 reveals the difference in F2 (Hertz) between TRAP and DRESS.

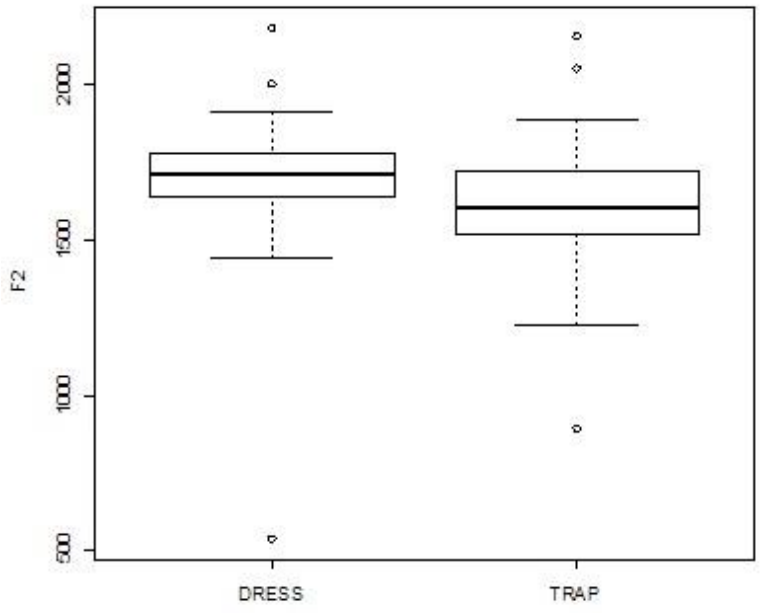

Figure 4. 18: Difference between DRESS (left) and TRAP in F2 (Hertz) of male low proficiency speakers

The situation is a bit different for male high proficiency speakers. Figures 4.19 and 4.20 demonstrate the distribution of English vowels produced by male high proficiency speakers in the two contexts.

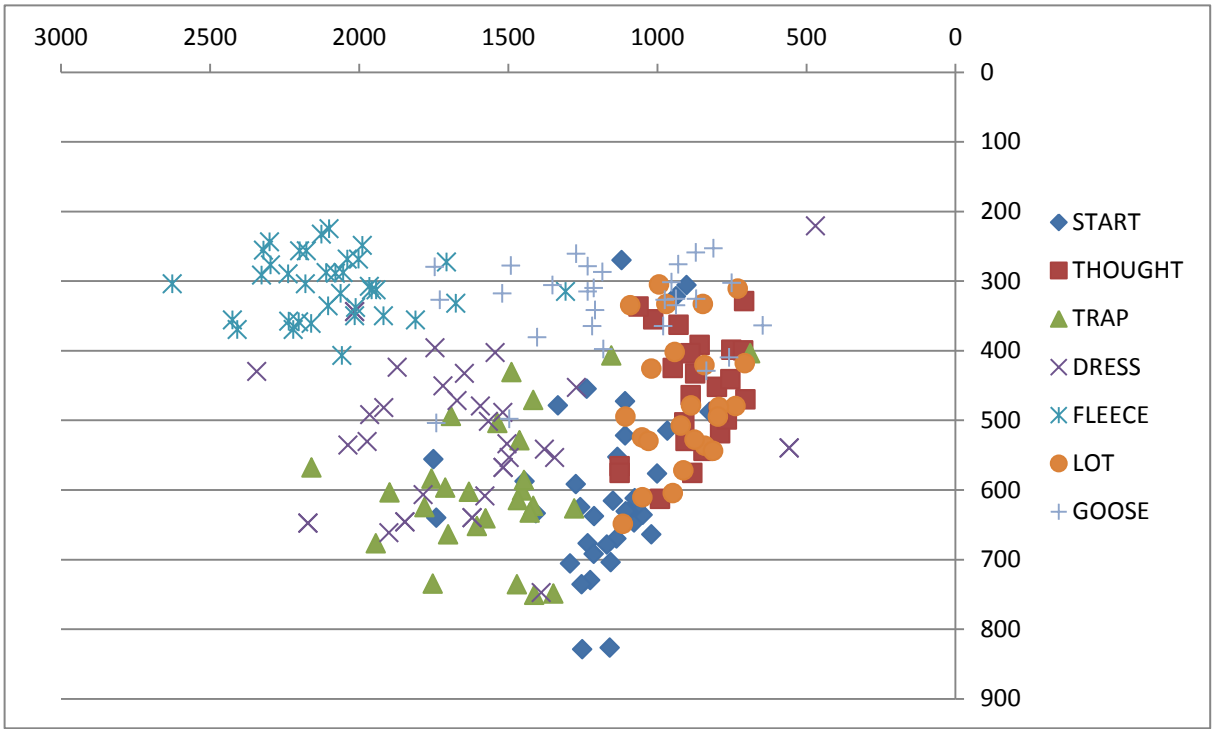

Figure 4. 19: English vowels produced by male high proficiency speakers in the Chinese context

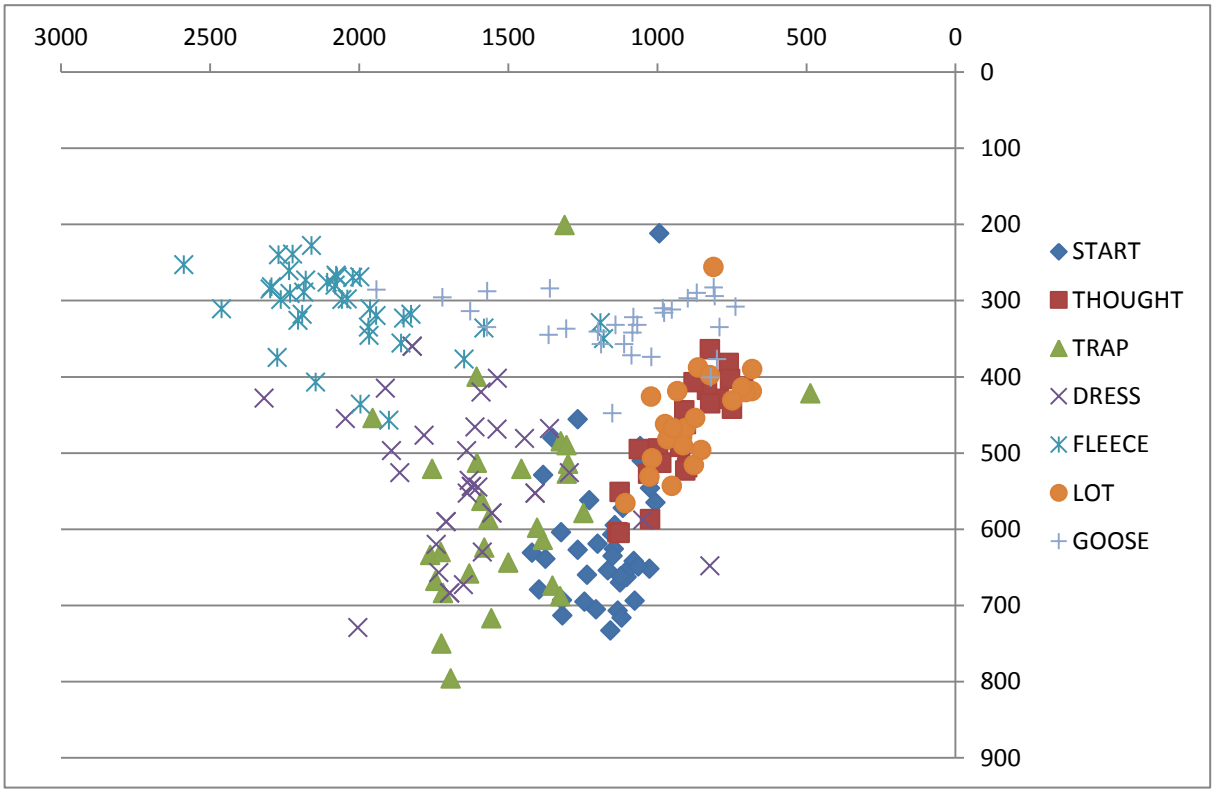

Figure 4. 20: English vowels produced by male high proficiency speakers in the English context
The results of Wilcoxon-tests are in Table 4.11.

	Chinese context		English context	
	F1	F2	F1	F2
TRAP and DRESS	0.01038	0.07656	0.09325	0.03248
LOT and THOUGHT	0.6876	0.2524	0.3172	0.4095

Table 4. 13: Results of p-value in Wilcoxon-test between TRAP and DRESS, and LOT and THOUGHT

The results for the difference between TRAP and DRESS are significant in F1 (p=0.01038) in the Chinese context and in F2 (p=0.03248) in the English context, and near significant in F2 in the Chinese context and in F1 in the English context. However, the difference between LOT and THOUGHT is not significant. The results above show that for the male high proficiency speakers, TRAP and DRESS are more separated but LOT and THOUGHT are still merged. Figure 4.21 shows the difference in F1 between DRESS and TRAP for male high proficiency speakers; whereas Figure 4.22 shows the difference in F2 between the two vowels for male high proficiency speakers.

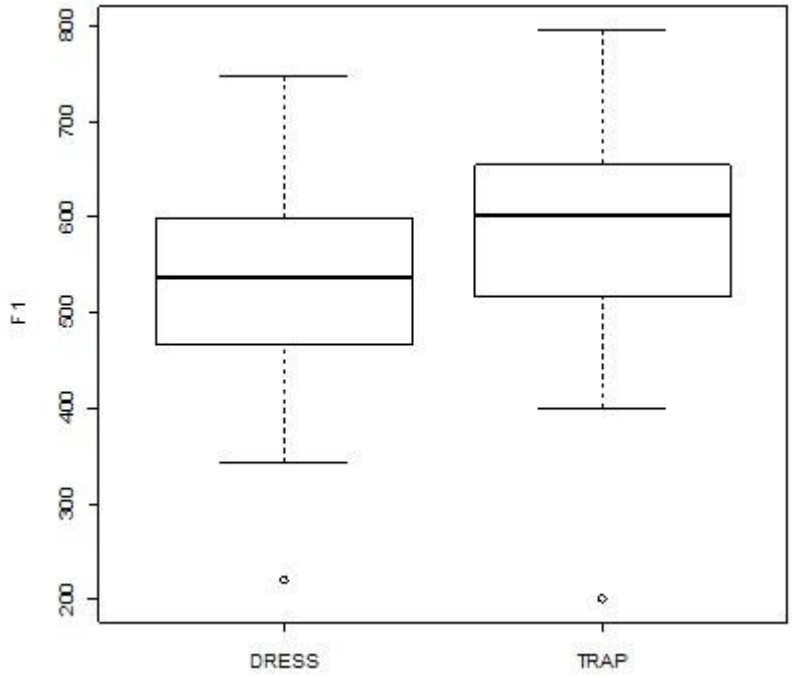

Figure 4. 21: Difference between DRESS (left) and TRAP in F1 (Hertz) for male high proficiency speakers

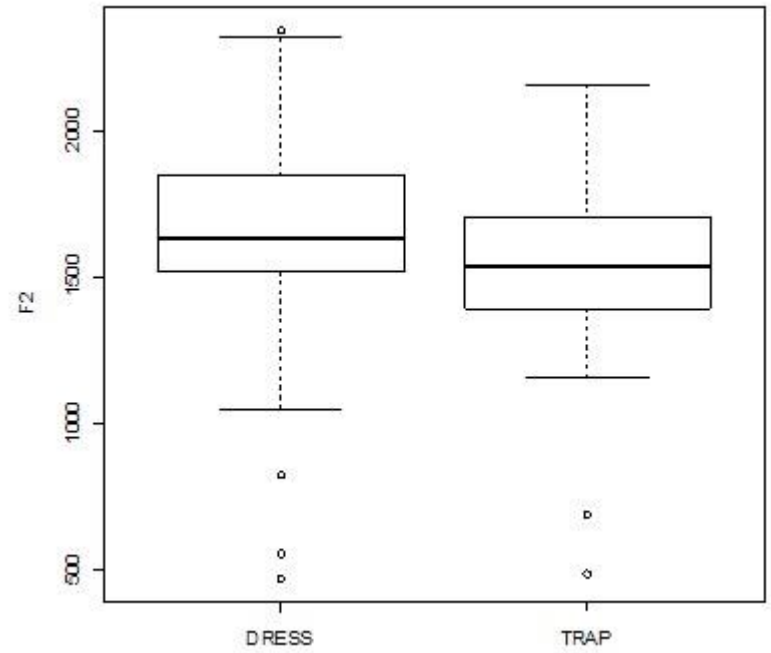

Figure 4. 22: Difference between DRESS (left) and TRAP in F2 (Hertz) for male high proficiency speakers

The results for females are not considered in these wilcoxon tests due to the lack of a sufficient number of female participants. But the female speakers are included in the analysis of Regression models for contextual effects, which uses mixed effects models to control for speaker variability.

Discussion

For Chinese /u/, results indicate that the English context causes this vowel to move toward its English counterpart, thus making it a more fronted /u/. The effect of the English context is not found on other vowels. Note that when the Chinese passage is read after the English context interview and the English passage, a change toward an English-like fronted GOOSE is a clear sign of the effect of short-term exposure to the English environment. High proficiency speakers can distinguish well between GOOSE and Chinese /u/. However, in terms of changes in Chinese as being affected by the English context, both high and low proficiency speakers reveal the effect.

When considering the three near-significant results of English vowels, two of them are only present for low proficiency speakers. Although the direction of the changes is opposite to our expectation, for vowels, low proficiency speakers obviously show more inter-lingual interference than high proficiency speakers. The incorrect direction of moving LOT, THOUGHT and START might be caused by the experimenter's identity. It is plausible that our Chinese bilinguals are not familiar with the situation of speaking English to a Chinese native speaker. Priming with the English context and the Chinese experimenter may have made them feel weird and reminded them of the experimenter's identity, thus eliciting more Chinese-like vowels. It has been shown by Hay, Drager and Warren (2009, 2010) that the identity of the experimenter can affect both speakers' speech production and perception. In our case, the effect of the experimenter's identity could influence production in the opposite direction of the context.

4.2.3: Summary of vowel results

Based on the results so far, we can see that the English context does indeed have an effect on speakers' Chinese vowel production. Additionally, high and low proficiency speakers behave differently in the English context. The relationship between productions of START and LOT and THOUGHT vowels suggests that it might have been inherited from the speaker's Chinese

production: because there is only one back centre vowel in Chinese. It is also important to remember that the trends in English production are only trends, whereas the effect on GOOSE reaches significance. Among all the vowels, only Chinese /u/ reveals a clear effect of L2 on L1, and this will be discussed further in the discussion chapter.

Chapter 5

Results of Consonants

Eight consonants are tested in the current research: /ʃ/, /ʂ/; /ʒ/, /r/; /tʃ/, /tʂ^{h}/; and /dʒ/, /tʂ/. The results are shown below.

5.1 English /ʒ/ and /r/, and Chinese /r/

The English fricative /ʒ/ normally occurs in the final syllable of a word, for example in 'leisure', 'treasure' and 'pleasure'. The Chinese approximant /r/ can be located in the C slot in a C-V structure. The pairing of the two sounds is based on the observable errors in English /ʒ/ produced by Chinese learners. This type of error is also mentioned by Wu (2008) as a common mistake made by learners who speak a Northern dialect.

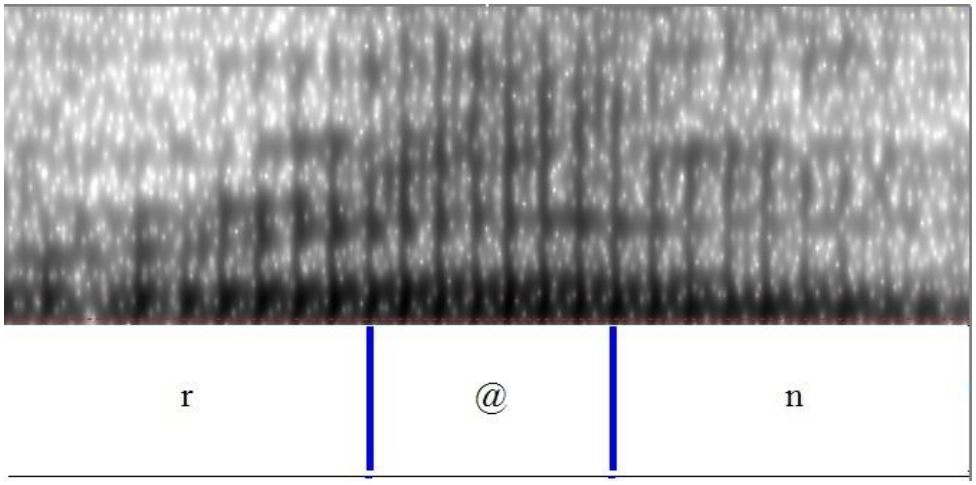

Figure 5. 1: Spectrogram of a the Chinese word 'jian ren'

Figure 5.1 shows that like English /r/, Chinese /r/ clearly shows its F1, F2 and F3, which indicates that like English /r/, Chinese /r/ is a rhotic consonant. All of the Chinese words for testing have a C-V-C-V (-C) structure. Acoustically the two consonants have few similarities. /ʒ/ is a fricative so a spectrum would display frication of high frequency, whereas /r/ is an approximant which would show its F1, F2 and F3 in its duration. However, in the experiment, some of the speakers failed to produce /ʒ/ but instead produced /r/ when speaking English. The two graphs in Figure 5.2 present an acoustical comparison between /ʒ/ and /r/. Both sounds occur in the English words produced by the same Chinese male speaker.

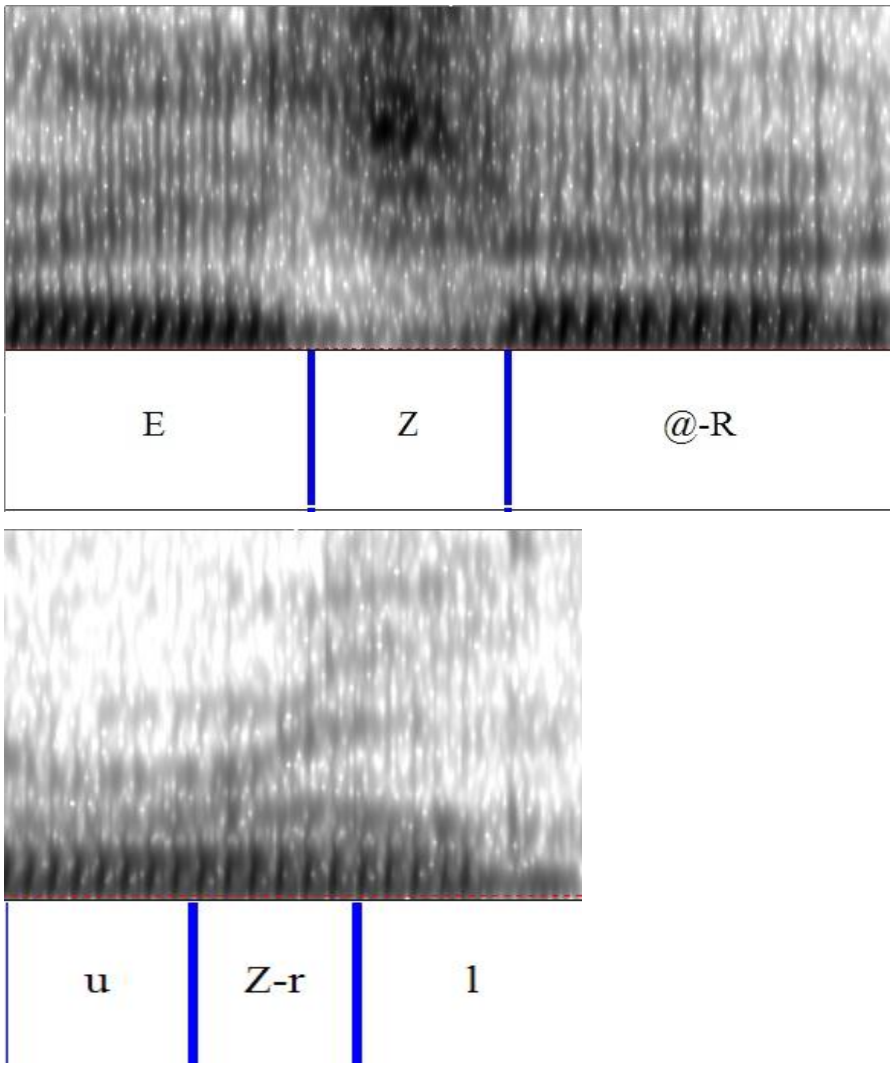

Figure 5. 2: A comparison between /ʒ/ produced in the word ‘pleasure’ (top) and /r/ produced in the word ‘usual’

The substitution of /ʒ/ with /r/ can be found in reading speech productions by both high and low proficiency English speakers in the experiment. Tested words ending in a schwa can trigger /r/-substitution. People who speak a Northern dialect, especially people from the area around Beijing, are familiar with /r/ sound. This is not to say that other dialects do not have an /r/ sound, but speakers of Northern dialect are more influenced by the Beijing dialect, which contains a large number of rhotic /r/ flavored vowels, which is a feature not found in other dialects. When speaking English, Northern dialect speakers also exhibit such a tendency of rhoticsing. For example, *usually* would become *urrually*. Among our tested tokens, only decision and vision do

not contain /r/ or /l/. Since thirteen participants are Northern Chinese, /r/-substitution occurs with a high frequency in the experiment. However, speakers never substitute /ʒ/ with /r/ in 'decision' and 'vision'. Perhaps, the substitution does not occur in words ending in '-sion'. For high proficiency speakers, their percentage of /r/-substitution is slightly lower in the English context than that in the Chinese context, but low proficiency speakers do not show this tendency. More importantly, for low proficiency speakers, the number of /r/-substitutions is 13 out of 56 in the Chinese context and 14 in the English context, which is much higher than high proficiency speakers' 4 and 2 respectively. There is no significant difference between high and low proficiency speakers in terms of context, as shown by a Fisher's exact test. Table 5.1 gives the number of /r/-substitution and count value in each context.

low	measure	usual	treasure	leisure	pleasure	casual	vision	decision	total
speaker1	CE	CE	E	CE	CE				9
speaker2		CE							2
speaker6		CE							2
speaker9		CE	CE	CE		CE			8
speaker10									0
speaker11									0
speaker14	CE	CE		CE					6
total	4	10	3	6	2	2	0	0	27
high	measure	usual	treasure	leisure	pleasure	casual	vision	decision	total
speaker3		C			C				2
speaker4									0
speaker5									0
speaker7		CE							2
speaker8									0
speaker12									0
speaker13		CE							2
total	0	5	0	0	1	0	0	0	6

Table 5. 1: Summary of r-substitution by both high and low proficiency speakers by word (speakers in grey colour represent Chinese Northern dialect)

The table shows the insertion of /r/ when there is /r/ or /l/ in the word. /r/-substitution did not occur when there was no /r/ or /l/ already in the word, such as *vision* and *decision*. /r/-substitution did not occur in the speech of all the northern dialect speakers. Given the fact that /ʒ/ does not

exist in Chinese production; statistical tests on the acoustics were carried out only between /r/-production in Chinese and /r/-substitution in English. We were interested in whether the /r/s being produced in English resembled the Chinese /r/ phonetically, and whether this differed across the English and Chinese context. As has been mentioned, each consonant has eight tokens, thus F3 value for /r/-production in Chinese are taken from eight tested words. In all of the words, /r/ is located at the beginning of the second syllable in a C-V (-C) - C-V (-C) structure. F3 values are taken as a measurement which can represent /r/-phoneme. The factor 'context' is tested as a potential predictor in speakers' Chinese and English productions.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	2188.71	2193.396	1851.2	2526.9	0.0001	0
Context= English	56.74	-1.284	-326	323.5	0.9974	0.6421

Table 5. 2: Result of a regression model with context as a fixed effect from speakers' /r/-substitution in English

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	2233.44	2233	2075.63	2399.7	0.0001	0
Context= English	99.82	100.2	15.86	191.6	0.0238	0.0249

Table 5. 3: Result of a regression model with context as a fixed effect from speakers' Chinese /r/ production

Table 5.2 gives a mixed effects regression model with speaker and word as random effects and context as a fixed effect from speakers' English production. Table 5.3 tests a mixed effects regression model with the same effects but from speakers' Chinese production. Interestingly, the result is significant in the Chinese production but not in the English production. A $p = 0.0249$ and estimate value = +99.82 indicate that when speaking Chinese, the English context will increase F3 of /r/. A comparison can be made between Chinese /r/ and English /r/ (substituted /r/ not included).

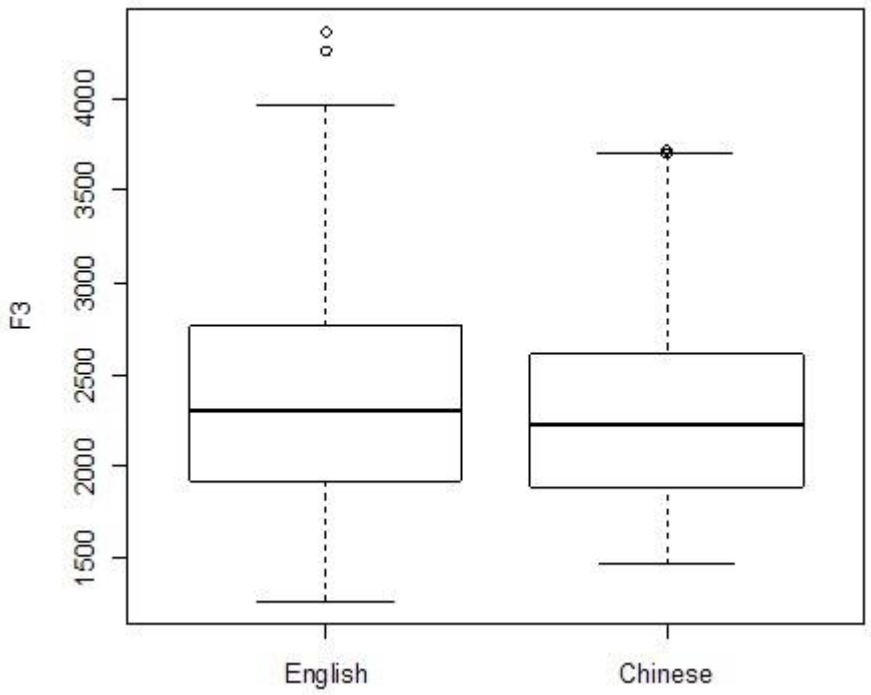

Figure 5. 3: Comparison of F3 between /r/ in Chinese and /r/ in English

Figure 5.3 gives a comparison of the two sounds. F3 of English /r/s are taken from real /r/. There were seven English tokens, all of which are two syllable words that have /r/ at syllable-initial position or part of a syllable-initial consonant cluster. Wilcoxon test shows a result of p-value = 0.02317, which means that F3 of English /r/ is significantly higher than that of Chinese /r/. Given the positive estimated value when speaking Chinese, higher F3 represents less constriction, which means a more English-like /r/.

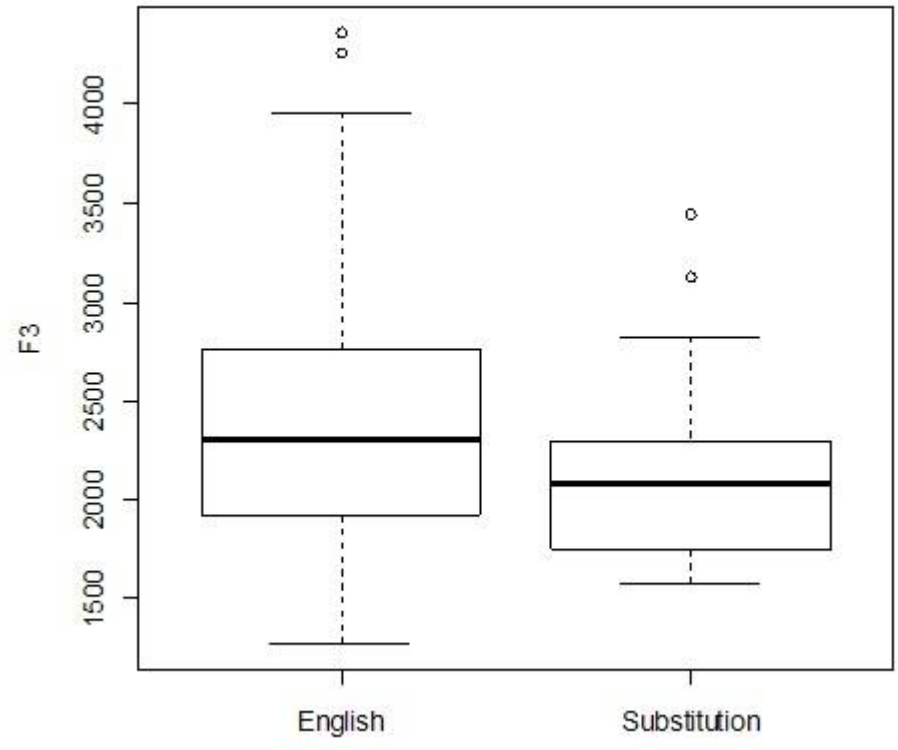

Figure 5. 4: Comparison of F3 between /r/ in English and /r/-substitution

Figure 5.4 shows comparison between substituted /r/ and real /r/ in English. A Wilcoxon test shows a result of p-value = 0.01298, which reveals that F3 of substituted /r/ is significantly lower than that of real /r/. When comparing Chinese /r/ with substituted /r/ in English, Wilcoxon test shows a result of p-value = 0.0795, suggesting that the two types of /r/ are not significantly different. A regression model was also conducted for real /r/ in English and the results are given in Table 5.4.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	2425	2420.1	2280.7	2547.02	0.0001	0
Context= English	-83	-83.34	-145.5	-21.4	0.0092	0.0074

Table 5. 4: Results of a regression model with context as a fixed effect from speakers' English /r/ production

The results suggest that speakers substituted in Chinese /r/ when speaking English and substituted for /ʒ/ - regardless of context. In Chinese, the English context leads to more 'English' /r/ production. But in English, the English context leads toward more 'Chinese' /r/ production.

The results of /r/ suggested that our Chinese speakers substituted Chinese /r/ for English. The substituted /r/ was not affected by context. However, /r/ production showed some effect of context that was not only in Chinese but also in English. The difference between English /r/ and Chinese /r/ suggested that our speakers perceived them as different categories.

5.2: Consonants with frication

As mentioned in the methodology, the current study explores the frequency at the peak and amplitude difference as possible measurement. To explore the feasibility of such a method, a comparison of the frequency and amplitude difference between English consonants and their Chinese counterparts was carried out.

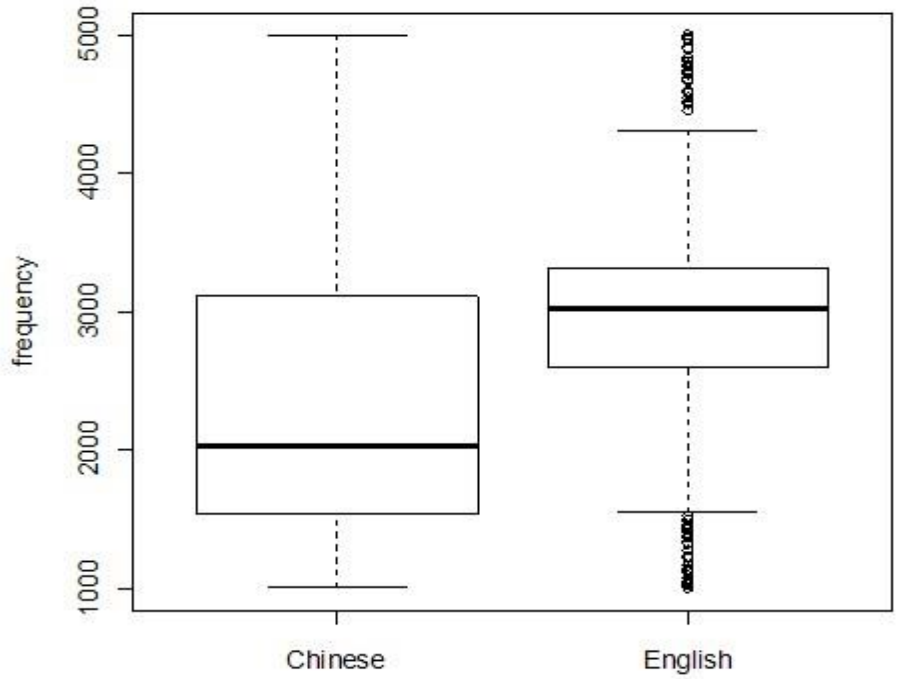

Figure 5. 5: Comparison of the frequency between /ʂ/ in Chinese and /ʃ/ in English

Figure 5.5 above shows comparison of the frequency at the peak between Chinese /ʂ/ and English /ʃ/. The result of a Wilcoxon test is p-value < 0.001, suggesting that the frequency in Chinese frication is significantly lower than that in English.

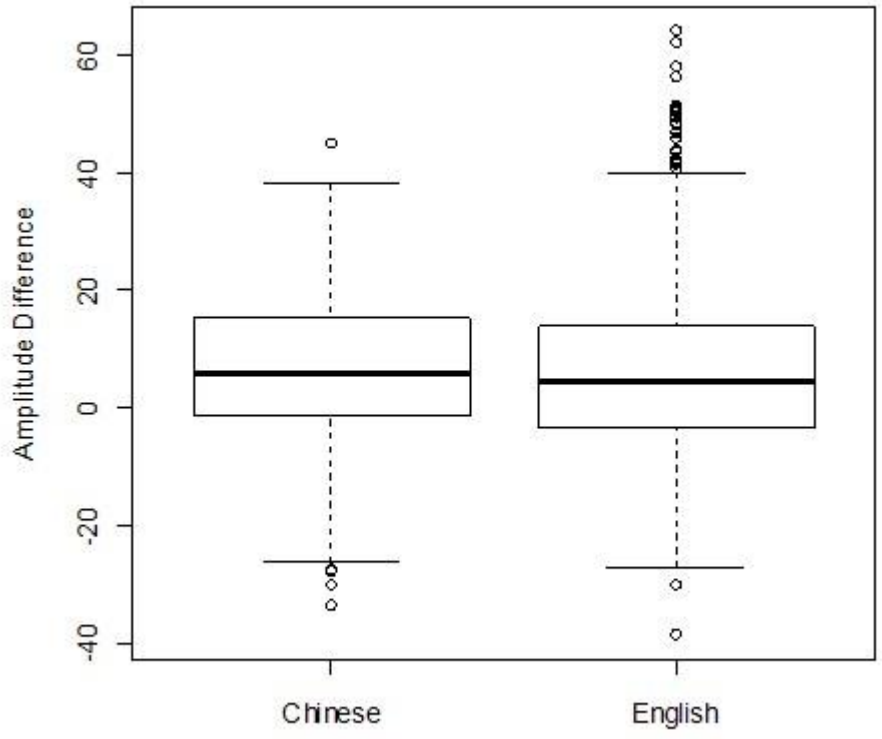

Figure 5. 6: Comparison of Amp Diff between /ʂ/ in Chinese and /ʃ/ in English

Figure 5.6 above shows the comparison of amplitude difference between Chinese /ʂ/ and English /ʃ/. As mention earlier, amplitude difference is the difference between the amplitude at zero frequency and amplitude at the highest peak. The result of a Wilcoxon test is p-value = 0.02806, which means that the Amplitude Difference in Chinese frication is significantly larger than that in English. The result is consistent with our initial observation. Given the fact that English /ʃ/ has some outliers with a high Amp Diff value, the difference of Amp Diff between the two sounds is not salient. However, a difference in the dimension of frequency between the two sounds can be seen in figure 5.7. The graph was created with two dimensional scales: the frequency as x-axis and amplitude difference as y-axis.

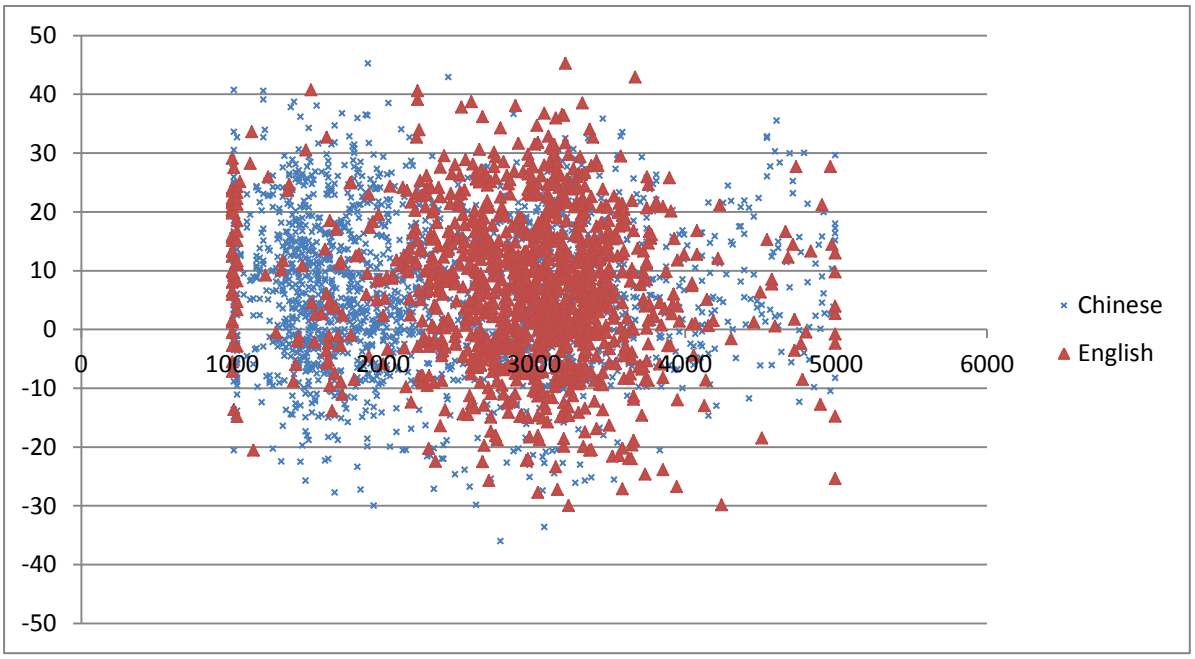

Figure 5. 7: Distribution of Chinese /ʂ/ and English /ʃ/

The graph shows the distribution of Chinese /ʂ/ and English /ʃ/. The frequencies of the two sounds almost distribute separately on the x-axis, whereas on y-axis, amplitude difference can hardly separate the two sounds.

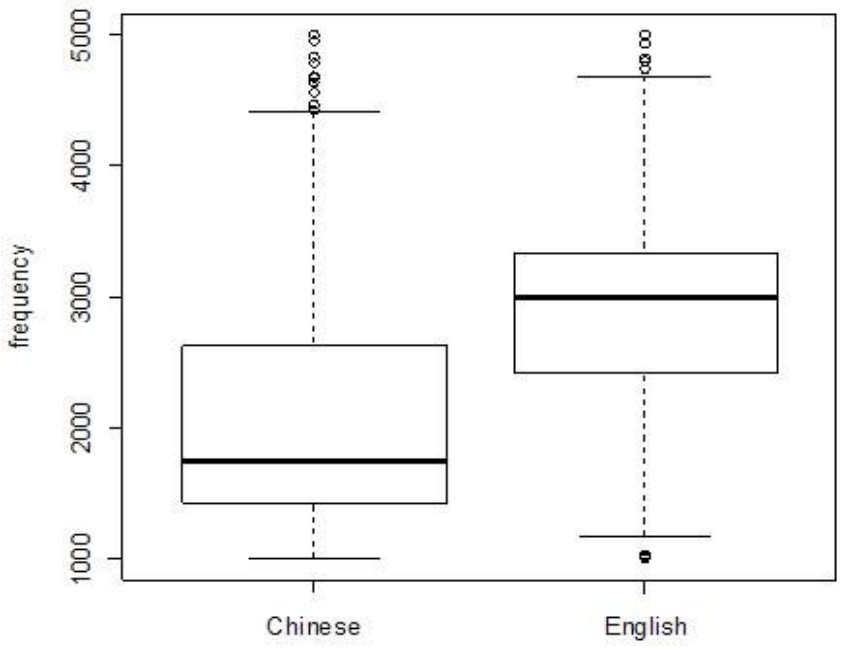

Figure 5. 8: Comparison of the frequency between English /tʃ/ and Chinese /tʂʰ/

Figure 5.8 shows the comparison of frequency at the peak between English /tʃ/ and Chinese /tʂʰ/. The result of Wilcoxon test shows p-value < 0.001, which indicates that the frequency in English /tʃ/ is significantly higher than that in Chinese /tʂʰ/.

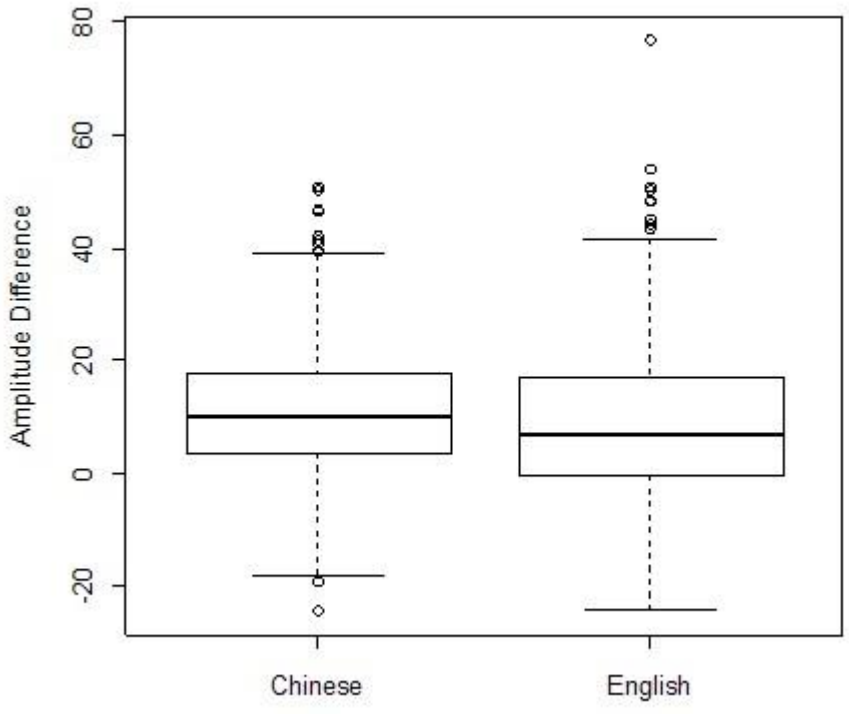

Figure 5. 9: Comparison of Amp Diff between English /tʃ/ and Chinese /tʂʰ/

Figure 5.9 gives the comparison of amplitude difference between English /tʃ/ and Chinese /tʂʰ/. The result of a Wilcoxon test gives a p-value = 0.0001931, suggesting that Amp Diff in English /tʃ/ is significantly smaller than that in Chinese /tʂʰ/, although again, this is not as marked as the frequency measure.

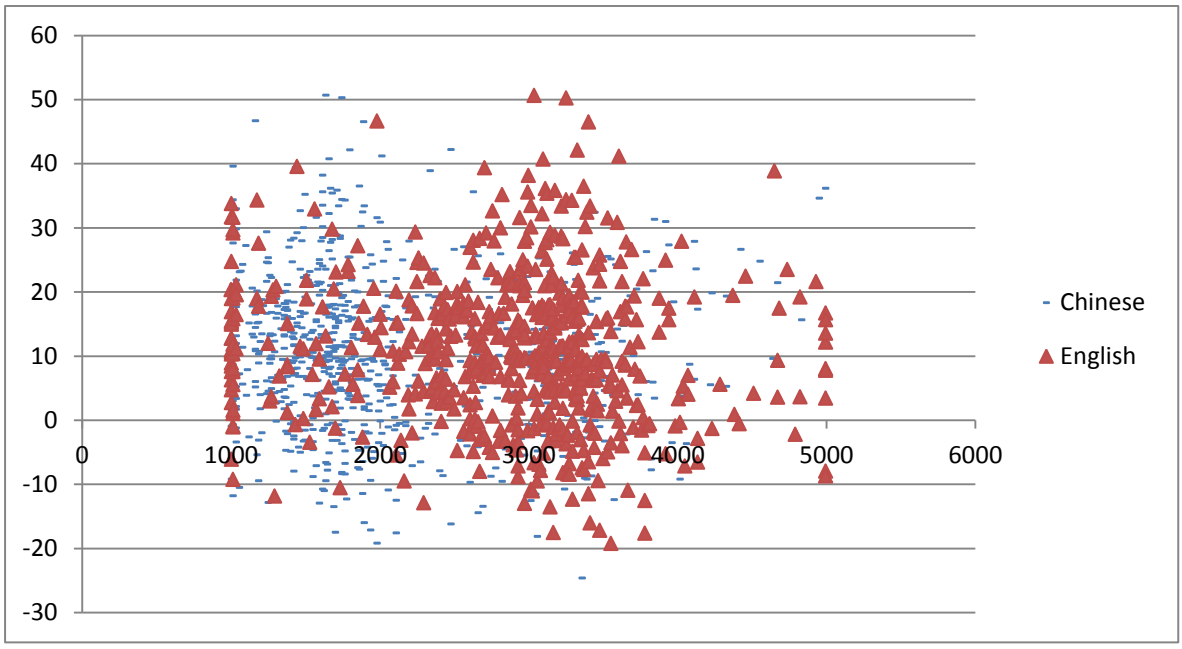

Figure 5. 10: Distribution of English /tʃ/ and Chinese /tʂʰ/

Similar to /ʃ/-/ʂ/, there are slight differences of Amp Diff between the two sounds. However, the difference of the frequency is salient enough to be seen. The graph above shows the distributions of the two sounds. The differences of the frequencies on the x-axis are large enough to distinguish the two sounds. However, on the y-axis the two sounds can hardly be separated.

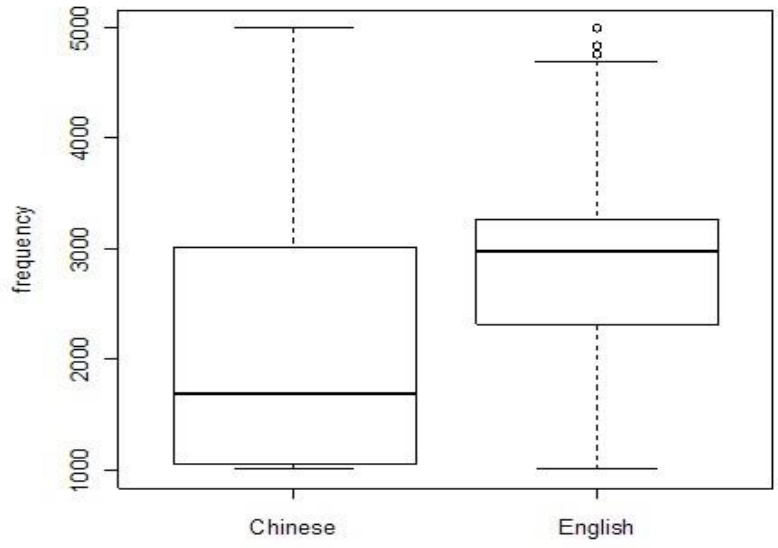

Figure 5. 11: Comparison of the frequency between English /dʒ/ and Chinese /tʂ/

Figure 5.11 shows comparison of the frequency at the peak between English /dʒ/ and Chinese /tʂ/. The result of a Wilcoxon test shows p-value < 0.001, suggesting that the frequency in English /dʒ/ is significantly higher than that in Chinese /tʂ/.

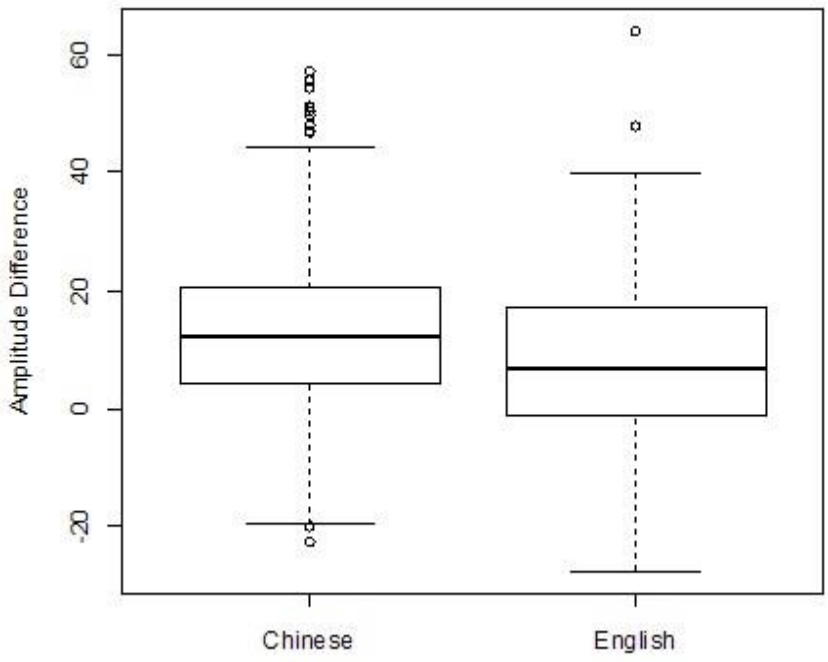

Figure 5. 12: Comparison of Amp Diff between English /dʒ/ and Chinese /tʂ/

Figure 5.12 shows the comparison of amplitude difference at the peak between English /dʒ/ and Chinese /tʂ/. The result of a Wilcoxon test shows a p-value = 1.040e-09, which means Amp Diff in English /dʒ/ is significantly smaller than that in Chinese /tʂ/. Again, the separation is not as extreme as that for frequency.

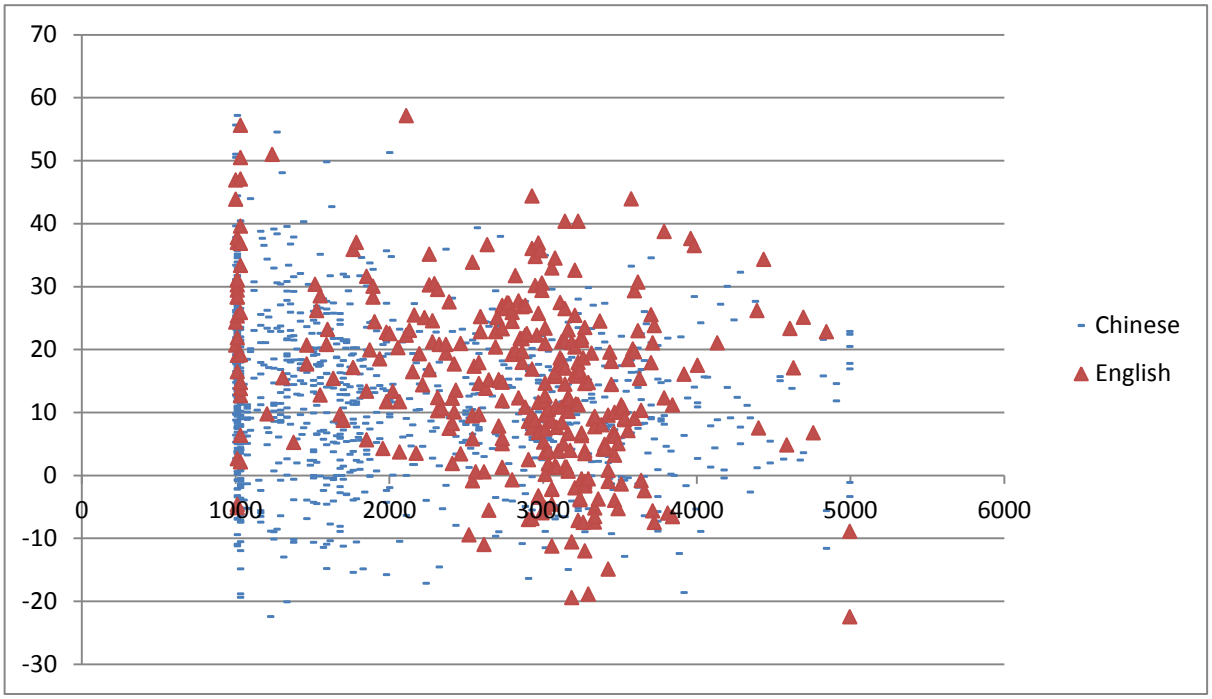

Figure 5. 13: Distribution of English /dʒ/ and Chinese /tʂ/

The distribution of /dʒ/-/tʂ/ is similar to the other two pairs of consonants. While there are slight differences in Amp Diff, the differences in the frequency are very salient.

The distributions of the three pairs of consonant share some similarities. On the one hand, the correlation between the frequency and amplitude difference is weak. On the other hand, although amplitude differences in English consonants are always smaller than those in Chinese, the differences are not large enough to very clearly distinguish a retroflex sound from a post-alveolar sound. In other words, although there are differences between English consonants and Chinese ones in terms of Amp Diff, the differences on the y-axis are not large enough to separate the distributions of the two languages. Based on the three box plots showing amplitude difference above we can see that although the differences in Amp Diff between English and Chinese consonants are significant, most of the pairs are overlapping. The frequency of the peak in Chinese sounds can range between 1000Hz to 5000HZ, but they mostly focus at around 2000Hz. Based on the comparison and distributions of consonants from the two languages, it can be concluded that the frequency at the peak is a clear measure to distinguish post-alveolar frication

from retroflex one. However, amplitude differences are only slightly smaller in English consonants, making it less reliable to function as a divider. It also suggests that Chinese consonants are more voiced. Based on this, a decision was made to concentrate on frequency as the dependent variable in the analysis of factors influencing the details of fricative production.

5.2.1: tested fricatives

Six consonants were tested and only the frequency at the peak was tested as a parameter. The six consonants were: English /ʃ/ and Chinese /ʂ/, English /tʃ/ and Chinese /tʂ^{h}/, and English /dʒ/ and Chinese /tʂ/. English /ʃ/ tokens in the experiment were all in word-initial position, for example, *shoebox* and *sugar*. In Chinese, /ʂ/ is located at either consonant position in a C-V-C-V structure such as *shufu* and *dushe*. The position of English /tʃ/ is either word-initially or at the end of a word. For example: *charger* and *touch*. Selected testing targets were all two-syllable words. The position of the Chinese consonant /tʂ^{h}/ is at either consonant position in a C-V-C-V structure such as *chabie* and *qingche.* The analyses only dealt with the frication part of the affricates. English /dʒ/ in our experiment is located in either word-initial or word-final position. For example, the testing targets include *large* and *joke*. The position of Chinese /tʂ/ is at either consonant position in a C-V-C-V structure such as *zheyan* and *guozhi*.

Regression models were tested for each consonant. The major tested factors were the interaction of context and proficiency, and gender, with speaker as random effect. The Praat script which originally extracted the spectra did not retain the word identity, so we were unfortunately unable to include word as random effect. If the interaction was not significant, then context and proficiency were tested as non-interacting factors. There would be two outcomes. If either was significant, then the other was dropped and to check whether it was still significant; if neither was significant, then the least significant one was dropped and to see whether the other one could become significant. Gender was always kept in the model, as a central factor.

English /ʃ/

Nothing was found significant for English /ʃ/ except that males produced /ʃ/ with lower frequency than females.

Chinese /ʂ/

Regression model showed that Chinese /ʂ/ was produced with significantly higher frequency in the English context. However, proficiency was not found to be significant. The results are given below.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	2528.6	2529.7	2140.72	2920.8	0.0001	0
Context=English	168.3	168.8	90.31	247.2	0.0002	0
Gender=M	-333.7	-336.4	-812.53	114.7	0.1432	0.1523

Table 5. 5: Results of regression model for testing Chinese /ʂ/

Given the fact that English frication has higher frequency than Chinese one, the positive estimated value of the English context shows that Chinese /ʂ/ is more English like in the English context.

English /tʃ/

For English /tʃ/, the interaction was found significant; which means that context indeed matters, but it affected low and high proficiencies speakers differently. The results are given in Table 5.6.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	2940.7	2941.4	2614.88	3273.734	0.0001	0
Context=English	144.9	142.8	-25.85	311.772	0.096	0.1009
Pro=low	128.7	125	-215.44	454.111	0.4382	0.4178
Gender=M	-251	-250.5	-582.22	70.348	0.1268	0.1089
Context=English: pro=low	-268	-263.7	-535.1	-4.188	0.052	0.0491

Table 5. 6: Results of regression model for testing English /tʃ/

For low proficiency speakers, the English context lowers the frequency of /tʃ/, whereas for high proficiency speakers, the English context increases the frequency of /tʃ/. To see whether one group was carrying the effect, low and high proficiency speakers were tested in separate regression models. Results are given in Tables 5.7 and 5.8.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	2869.3	2870.5	2433.19	3301.3	0.0001	0
Context=English	144.8	145	-19.92	316.4	0.09	0.0904
Gender=M	-155.4	-156.2	-643.26	324.1	0.4708	0.4442

Table 5. 7: Results of regression model for testing English /tʃ/ for high proficiency speakers

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	3141.9	3138.3	2523.3	3681.9	0.0001	0
Context=English	-123.7	-125.9	-338	87.8	0.2472	0.2531
Gender=M	-359.6	-357.3	-1068.7	338.9	0.2434	0.1683

Table 5. 8: Results of regression model for testing English /tʃ/ for low proficiency speakers

The two tables above indicate a difference between high and low proficiency speakers. The English context affects high proficiency speakers nearly significantly, and a positive estimated value suggests that the English context increases the frequency (i.e. invokes a more English-like fricative). However, nothing was found significant for low proficiency speakers, and there was a trend of the English context to decrease the frequency for them.

Chinese /$t_{ʂ}^{h}$/

Regression model showed that Chinese /$t_{ʂ}^{h}$/ was produced with higher frequency in the English context. However, proficiency was not found significant. The results are given in Table 5.9.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	2049.4	2049.8	1694.25	2406.4	0.0001	0
Context=English	158.4	158.2	54.68	272.1	0.0032	0.0045

Gender=M	-140.6	-141.7	-570.2	292.5	0.4868	0.5108

Table 5. 9: Results of regression model for testing Chinese /tʂʰ/

Given the fact that English frication has higher frequency than Chinese one, positive estimated value of the English context shows that Chinese /tʂʰ/ is more English like in the English context.

English /dʒ/

Nothing was found significant for English /dʒ/: neither context nor proficiency had an effect on the sound.

Chinese /tʂ/

Regression model shows that Chinese /tʂ/ was produced with higher frequency in the English context. However, proficiency was not found significant. The results are given in Table 5.10.

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(>|t|)
(Intercept)	2336.3	2337.4	1969.61	2718.274	0.0001	0
Context=English	161.5	161.6	42.25	278.064	0.0086	0.0079
Gender=M	-432.8	-436.2	-869.22	8.889	0.0504	0.0417

Table 5. 10: Results of regression model for testing Chinese /tʂ/

Given the fact that English frication has higher frequency than Chinese one, the positive estimated value of the English context shows that Chinese /tʂ/ was more English like in the English context. Gender was also found significant for this sound: males produced the sound with lower frequencies than females.

5.2.2: Summary

Based on the above results, we can conclude that proficiency generally does not affect consonant realisation. Every tested Chinese consonant had a higher frequency when produced in the English context, revealing the effect of the English context on Chinese. The effect of context is stronger in Chinese than in English. For English /tʃ/, high proficiency speakers showed near significant effect of context, which is not found for low proficiency speakers.

Chapter 6

Discussion of Results

As mentioned in the introduction, there are two major research questions arising from the current study. Firstly, the effect of context on languages across different speakers with different proficiencies could help explore the internal relationship between two languages for late life bilinguals. Second, the phonetic drift in the language used by the individual participants could be affected by the short-term recent exposure to the context of another language.

6.1: Transfer effects

Transfer effects can be observed from results in the current study. For example, the distinction between TRAP and DRESS is clearer for high proficiency speakers than for low proficiency speakers. However, speakers from both groups had difficulty in separating LOT from THOUGHT very well. It is not surprising that neither group could differentiate LOT from THOUGHT, because the lack of a competitive central-back vowel to Chinese /o/ results in a merged category for THOUGHT and LOT for the bilinguals. In addition, the distance from START to LOT and THOUGHT suggests that it might be transferred from the difference between Chinese /a/ and /o/.

6.2: Contextual effects

	effect of English context on Chinese	effect of Chinese context on English
/ʃ/-/ʂ/	●-s→E B	○
/tʃ/-/tʂ^{h}/	●-s→E B	●-t→C H
/dʒ/-/tʂ/	●-s→E B	○
START-/a/	○	●-t→E L
LOT-/o/	○	●-t→E L
THOUGHT-/o/	○	●-t→E B
GOOSE-/u/	●-s→E B	○
FLEECE-/i/	○	○
DRESS-/e/	○	○
TRAP-/e/	○	○

substituted /r/	○	
English /r/ - Chinese /r/	●-s→E B	●-s→E B

Table 6. 1: Overall results of context effect on the two languages (●-effect of context; ○- no effect of context; s-significant; t-trend; →E-direction to English; →C-direction to Chinese; B-both proficiency; H-high proficiency; L-low proficiency)

Table 6.1 illustrates the overall effect of context on the two languages. Both vowels and consonants can be affected by context. Note that when some contextual effects are significant to some segments, high and low proficiency speakers do not perform significantly differently. Furthermore, all the significant results are found in Chinese rather than English. Additionally, there are trends that the Chinese context makes LOT and START become more English-like for low proficiency speakers and makes THOUGHT become more English-like for both high and low proficiency speakers. The Chinese context also makes English /r/ become more English-like for both proficiency groups.

The only predicted result of vowels was found with Chinese /u/, which is affected by the English context and moves toward English GOOSE. F2 of Chinese /u/ increases by 5.25% in the English context for both proficiency speakers. These findings prove that the effect of L2 on L1 can be triggered by very recent language exposure. The current experiment uses fricatives and affricates as testing targets, which are mostly distinguished by comparing post-alveolar sound to retroflex sound. More evidence is found by analysing four pairs of consonants. Part of the result of consonant production provides further evidence of the effect of the English context on Chinese sound production. By analysing English /r/ and Chinese /r/, it is found that the English context can affect Chinese approximant /r/ to move toward a less rhotic direction. However, English /r/ moves to a less rhotic direction in the Chinese context.

The analyses of three pairs of post-alveolar and retroflex sound provide more information. Comparison of the two types of sound showed that of the measures considered, only peak frequency can be used as a reliable measure. The English context is found to increase the frequency of all three retroflex sounds, thus making them post-alveolar-like. When the English context significantly affects the Chinese retroflex sounds, there is no proficiency affecting any of the three pairs. However, when the English context can increase the frequency of English /tʃ/,

proficiency has slight effects on the sound. The effect of the English context on the sound is found near significant for high proficiency speakers but not for low proficiency speakers.

6.3: Combining the two types of effects

Phonetic transfer effects and contextual effects can be used as a combination to explore the connection between L2 acquisition and priming effects. Assuming that sound changes, which were affected by the language context, result from the experience of the environment of that language, then a short-term exposure activates the language environment. Significant contextual effects in the present study can be seen as a short term phonetic transfer from L2 (English) to L1 (Chinese). The current study indicates that late bilinguals can build new phonetic categories for certain sounds. The SLM (Flege, 1987) states that if a speaker does not perceive a L2 sound as 'new', it will be perceived as 'similar' and included in the existing phonetic category. In this study, significant results suggest that each pair of sounds in the six pairs (five pairs of consonants and one pair of vowels) should belong to different categories for the participants. In other words, category formation occurs to these sounds. However, if new categories are formed, one might ask why contextual effects could be found at all, because based on the SLM, sounds sharing common space should have inter-lingual interference.

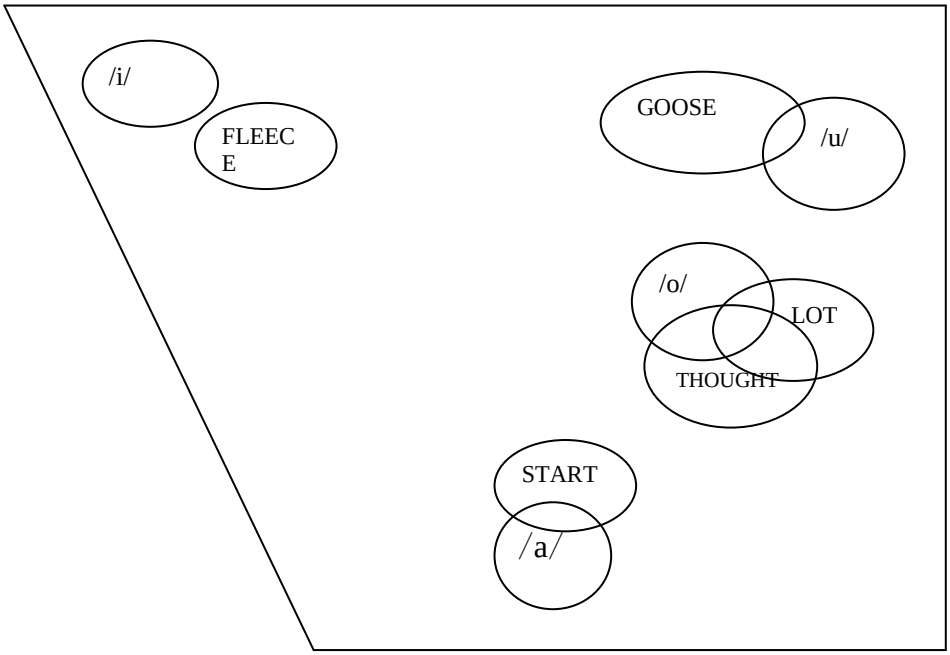

Figure 6. 1: A concept of category formation for all the participants. Lexical set words indicate English categories. Chinese categories are shown using phonemic representations.
Results suggest that the participants have formed new categories, but are not completely separated from existing L1 categories. Figure 6.1 illustrates this concept.

The idea of Figure 6.1 is consistent with Flege's SLM, which predicted the dissimilation of a sound from its neighbour sound when a new category is found for the sound. The dissimilation often appears as a continuum. Figure 6.1 suggests that for the absence of interference within a pair of vowels, an L2 sound is completely dissimilated from its L1 counterpart. For example, GOOSE-/u/, the two sounds are significantly different, but have still merged space. In addition, the formation of categories for some sounds could be easier or faster than the formation for other sounds. For example, the participants formed a new category for FLEECE, but not for THOUGHT. When priming with a language context occurs, phonetic elements in the overlapped area can influence each other, such as GOOSE-/u/, THOUGHT-/o/, LOT-/o/ and START-/a/. However, priming with a language context has no effects when the two phonetic categories are completely separated: e.g. FLEECE-/i/.

An interesting result is found between English real /r/ and Chinese /r/. The two sounds are significantly different: the Chinese one is more /r/-ful. This raises the question as to whether the Chinese /r/ is more rhotic. Future research needs to be done when data from Chinese monolingual speakers are collected.

Short-term priming with a different language is suggested to have an effect on the other language. The participants can shift their L1 into a style of L2 to accommodate the environment. In addition, high and low proficiency speakers did not perform differently when they shifted their L1 style. However, English sound moved to unexpected directions, which suggests that they might be triggered by different factors when speaking English.

In terms of the effect of L2 on L1, high and low proficiency speakers did not perform differently in this study. However, they did perform differently when speaking English. One interesting result is that some English phonemes seem to become somewhat more Chinese-like in the English context (although it is important to note that these effects were all borderline significant). One plausible explanation is that it could be caused by the experimenter's identity. The experimenter is late bilingual in Chinese and English, and the identity has the potential to influence the participants' production of sound. In the English context, the fact that English is spoken by a Chinese-English bilingual could appear marked, and thus reinforces the Chinese identity, which affects the following English reading and makes the English more Chinese-like. For low proficiency speakers, LOT and START were especially affected in this way but not for the high proficiency speakers.

This of course raises the question why having English spoken does not make the *Chinese* more Chinese like. Observations from this study suggest that when speakers are speaking Chinese, their identity as Chinese is being well expressed by the language choice. In this case, the exposure to English primes toward English phonetic realizations. High proficiency speakers are less likely to be affected by the experimenter's identity: their English /tʃ/ is slightly affected by the English context and moves to the more English-like direction. In the Chinese context, speakers did not need to express their identities because they were speaking their native language; as a result, it has no effect on their following Chinese reading.

For this reason, high proficiency speakers were less likely to be affected by the English-speaking Chinese experimenter than low proficiency speakers. It could be because high proficiency speakers are familiar with or have more experience in talking to a Chinese speaker in English, but low proficiency speakers may feel uncomfortable in conducting a conversation in this way. For future work, the same experiment could be conducted by a Pākehā English-Chinese bilingual, and research could be designed to test if the Chinese context could make the following Chinese reading become more English-like in that case. Alternatively, speakers can also be interviewed by two different interviewers: one Chinese and one Pākehā.

The participating speakers' overall English production might also be affected by the experimenter. The experimenter was a late bilingual speaker, who can distinguish between TRAP and DRESS, but not LOT and THOUGHT. Since high and low proficiency speakers were treated equally in the experiment, if the experimenter's inability to produce LOT and THOUGHT differently led the participating speakers to unable to produce the two sound differently, they could still be able to distinguish between TRAP and DRESS. However, low proficiency speakers' TRAP and DRESS were still somewhat merged. We would claim that the experimenter's English production was likely to have had minimal effects on the participants.

Another mysterious result in the current study was that despite the fact that contextual effect of L2 on L1 was found on the phonetic level, effects of L1 on L2 was almost not found. A possible reason could be that people have greater malleability in the native phonetic categories than in their L2. Material reading in L2 is usually very formal. Reading in L2 is so formal that it can hardly represent casual speech. In our case, L2 was so formally rehearsed; it was possible that there was no real malleability in the phonetic categories when speakers paid a large amount of attention to the reading speech. Some studies have already found that material reading contains fewer incorrect transfers and fewer errors in producing certain L2 segments (Dickerson & Dickerson, 1977; Dickerson, 1977 and Wenk 1979 as cited in Munro & Derwing, 1994). A so-called 'Chameleon Model' stated by Tarone (1979, 1982, 1983) suggests that inter-language as a continuum state is determined by how much attention is paid to the speech. When less attention is paid, more inaccurate sound will be produced, and therefore L2 production will be more accurate when more attention (more formality) is paid. In our case, material reading reflects formality, and thus producing less L1-L2 transfer and less inaccurate production are reasonable.

Consonants were more affected than vowels in the current experiment. All of the Chinese consonants were affected by the English context but only Chinese /u/ was affected. Note that even though the differences in frequency between Chinese and English fricatives were significant, the two types still cannot be completely separated on the dimension of frequency. In the current

study, articulation moving from post-alveolar to retroflex could be detected by measuring the frequency of frication. However this is not the case for vowels. Comparison between Chinese and New Zealand English vowels suggests that the pair of GOOSE-/u/ has the greatest distinction, thus the moving of /u/ can be detectable. Table 6.2 gives the calculated Euclidean distance between English sounds and their counterparts in Chinese (based on mean F1 and F2 values in this study).

vowels	Euclidean distance
/u/-GOOSE	139
/o/-THOUGHT	84
/o/-LOT	80
/a/-START	113
/i/-FLEECE	84
/e/-TRAP	65
/e/-DRESS	130

Table 6. 2: Euclidean distance between Chinese sounds and their English counterparts

On the contrary, distinctions between vowels within other pairs are not so large. Even if there are slight changes of vowels caused by the English context, they are still likely to be swallowed up by the large range of the variety.

There might be possible effects of the Chinese questionnaire on its following Chinese reading. The questions in the Chinese questionnaire were more oriented toward language. Language-related questionnaire might make our subjects pay more attention to their pronunciation. Based on this assumption, we would see ‘less formal’ speech production in the English context. Normally casual speech would cause contraction of the vowel space (Bradlow, Torretta & Pisoni, 1996). However, we did not see the contraction of the vowel space except for /u/, which indicated that speakers’ hyper-careful speech might not occur. The fact that we only see shifting of one vowel in Chinese, and that it is the vowel most distinct from its English counterpart, suggests that this is due to language priming and not to a change in formality.

On the other hand, the formality based explanation could potentially explain the results of English vowels. Formal speech enlarged the vowel space in the Chinese context, as a result, F1 of both LOT and THOUGHT were decreased. 'Careful speech' might also be able to explain the consonants results for the fricatives. If the participants performed 'more formal' speech in the Chinese context, Chinese fricatives would show more retroflexion, which were our results of the fricatives and affricates. However, 'careful speech' cannot give an explanation for the /r/-result in English. Chinese /r/s have lower F3. A lower F3 shows increased constriction, and by this argument we would expect lowered F3 in 'language' oriented, i.e. the Chinese context for both languages. However, we found that F3 of Chinese /r/ increased in the Chinese context, which contradicted the 'careful speech' argument. Such a fact also suggests that it is due to language priming but not formality.

6.4: Exemplar models vs. Abstractionist models

One controversial topic in terms of mental lexicon is how words are stored. A traditional abstractionist view suggests that words are stored as separate abstract phonemic categories (Pierrehumbert, 2016). Under this view, episodic information such as speech contexts, social information such as age and gender, and specific acoustic details are reduced to abstract phonemes that are phonological entities. In abstractionist models, producing a word involves activating a lexical representation of the word and phonological rules that govern phonetic representations into acoustic production (Ernestus, 2014).

On the other hand, exemplar-based models suggest that encountered utterances are stored with acoustic, social and contextual information (Johnson, 1997; Pierrehumbert, 2003). Exemplar theory suggests that mental representations are stored as exemplars that contain encountered instances. In these models, mental lexicon contains these exemplars and the exemplars represent different instances of tokens encountered. What have been reduced in the abstractionist view, such as social and contextual information and acoustic details, are also stored with exemplars. Based on exemplar models, speech perception is about matching the acoustic signal to the stored distribution of encountered utterances; whereas speech production is about to choose an average instance from the

distribution. Existing work has found that contextual details, voice, social information such as social class, age and gender, are also stored in our memories and therefore play an important role in speech perception and production (e.g., Sanchez, Hay & Nilson, 2015).

Apart from models that are based on a single view, recent studies suggest hybrid models of abstractionist models and exemplar-based models (e.g, Goldinger, 2007; Hawkins, 2003; McLennan, Luce & Charles-Luce, 2003; Pierrehumbert, 2002; 2016). For example, Goldinger (2007) explained a complementary model of abstract representation and episodic information in bilingual speakers. In this model, speech signals are inputs that form cortical representations. These representations are abstract. Episodic information about the speech, including specific characteristics about the speakers such as age, gender, social class, and acoustic information, is stored in the hippocampus that are responsible for managing memories. Information stored in the hippocampus is also sent back to the cortex and update new phonological rules. This model can also be applied to the perception-production in bilingual speakers. Based on Goldinger's (2007) complementary system, for L2 speakers, since cortical representations contain abstract phonological rules, acoustic information as input is filtered by L1 phonology before it is stored in the hippocampus, which means that L2 speakers' stored acoustic details could be different from those from native speakers of L2 (Maye, 2007).

Results from the current study can also be explained by a hybrid model of abstractionist models and exemplar-based models. Under such models, bilingual speakers' mental lexicon contains instances of words that are from both their L1 and L2. Not only does producing a word in one language activate exemplars at a word level, exemplars are also activated at a phonemic (abstract) level. Based on the SLM, sounds from two languages are perceived as a 'similar' sound, which can be explained by the effect of L1 phonology on L2 phonemic categorization (Maye, 2007). Activation of a sound in the merged category would also activate other 'similar' sounds in the same category. As a result, a 'similar' sound from the category was produced when producing a word in the other language. Results from the current study support such a hybrid model of exemplar theory and abstractionist view, suggesting that both abstract categories and episodic information are stored in bilinguals' mental lexicon.

6.5: Language Mode Framework

It is well documented that there are interferences between bilinguals' two languages. Grosjean (2012) argued for two types of interferences: static interference and dynamic interference. Static interference refers to the constant producing features from one language when speaking the other, whereas dynamic interference is temporarily producing features from the other language by accident. Simonet (2014) also discussed the distinction between the two types of interference, and used the term 'long-term interference' and 'transient interference' instead. Existing work that has been done earlier in the 1980s and 1990s mostly tested long-term interference (e.g., Flege, 1988; Flege, Munro & MacKay, 1995). In these studies, participants with different ages of learning, lengths of residence and quantities of language input were tested, which in order to test these factors. However, later work started to test transient interference. For example, Antoniou et al. (2011) tested the production of certain sound when bilinguals were code-switching, and found that code-switched sound in L2 were affected by the global language context of L1. Both Olson (2013) and Simonet (2014) tested the effect of immediate language context on sound production and found similar results: the effect of language context of L2 on the production of L1.

Interferences between bilingual speakers' two languages can be explained by the effect of language mode. According to Language Mode Framework (Grosjean, 2008), a bilingual speaker is constantly under the effect of the activation of two languages. The activation of two languages ranges from a unilingual mode, where only one language is activated, to a true bilingual mode, where the speaker's two languages are equally activated. There is a continuum between the two modes, and the degree of activating either language depends on the needs of the task, such as the language used by the interlocutor. Grosjean (2008) also argued that a true unilingual model is never achievable during experimental conditions, which means that one language is always more or less activated although a bilingual speaker is only using the other one.

Such a framework can also explain the current results. The global language is always activated when the bilinguals are reading in the other language, despite the fact that the global language is not being used. The effect of the global context is consistent with the effect of language context during

code-switching (e.g., Antoniou et al., 2011, Olson, 2013). As a result, sounds that are perceived as 'similar' between two languages were more likely to be affected by the very first language context.

6.6: Articulatory Setting Theory

The other alternative explanation of the current results is articulatory setting theory. It is believed that different languages have their own default articulatory gesture that might have effects when people are speaking a second language (Wilson & Gick, 2014). Evidence has been found from studies such as Gick et al. (2005) and Wilson, Horiguchi and Gick (2007). Wilson and Gick (2014) found that bilinguals who were perceived as native speakers showed distinct articulatory gestures, but those who are not perceived as native speakers did not show such differences. If languages are different in terms of underlying articulatory gestures, then there certainly would be interferences between two languages if certain couples of sounds from the two languages are produced in similar manners. For example, the Mandarin retroflex fricative is often produced in a similar way as its post-alveolar counterparts in English: the tongue is only slightly backward when producing the former than the later.

Chapter Seven

Conclusion

The aim of the current study has two focuses. The first was to discover the influence between L1 and L2 on late bilinguals, and to examine whether this was influenced by proficiency. Speakers were divided into two groups according to their use of English, which was based on their answers to the two questionnaires, and different proficiencies reflected the variance within the L1-dominant status. The second focus was to investigate short-term language priming effects on the two proficiency groups. To do this, speakers were recorded in two sessions – an 'English-dominant' session, and a 'Chinese-dominant' session. To achieve the different language primes, questionnaires were designed in two different languages with different questions and the experimenter needed to speak the priming language, which was same to the language of the questionnaire. This was followed by material reading first in the same language, and finally, in the non-primed language. Readings of the passages across sessions were compared.

We found that both vowels and consonants can be affected by different contexts. One of the Chinese vowels and all the four tested Chinese consonants move toward the direction of contextual language. There are also three English vowels and two English consonants moving toward the opposite direction of contextual language.

The experiment provided evidence for the first research question. We found the influence between L1 and L2 from late bilinguals. In terms of the effects between L1 and L2, low proficiency speakers were found to have more transfer effects than high proficiency speakers. This suggests that L1-L2 transfer is correlated to L2 proficiency. However, the two proficiency groups did not perform quite differently with respect to the effects of L2 on L1. Low proficiency speakers also showed more interference if near significant tendency could be considered. All these suggest that more proficient bilinguals might be better at suppressing the other language. Salvatierra and Rosselli (2011) in fact conducted a research on inhibitory control of bilingual speakers by comparing simple task and complex Simon task performed by young and elder bilinguals, also between bilingual and monolingual speakers. Their results suggest that bilinguals are better at inhibitory control than monolingual speakers, and bilingual advantage is correlated to age. Our results suggest that highly competent bilingual (relatively in the current study) might

be good at suppressing the other language, so that it does not interfere. A less practiced bilingual might not be good at this.

For the second research aim, priming effects were found between the two languages. The current study shows that within L1-dominant late bilinguals, inter-language interference occurs across a range of speakers' competencies. Speakers with different proficiencies do not perform differently when speaking Chinese priming with the English context. We may need to ask, if all of our L1-dominant speakers had not received enough L2 exposure/experience, and led both groups to the same results – i.e. perhaps the range of proficiencies was not wide enough to reveal any potential effect. However, unexpected results, possibly caused by the experimenter identity suggest that speakers with various degrees of English proficiencies might respond to a 'weird' Chinese English experimenter differently. Given the fact that high proficiency speakers conduct relatively higher percentage of English talking than low proficiency speakers, high proficiency speakers might have more experience of speaking English to a Chinese bilingual.

Overall, present study can contribute to the literature of bilinguals by providing evidence of inter-lingual interference between L1 and L2 phonetic systems amongst late bilingual speakers. Transfer effects suggest that inhibitory control might be correlated to L2 proficiency, so high proficiency speakers would have less interference between the two phonetic systems than low proficiency speakers. Results of priming effects suggest that L1-dominant speakers' language mode shifting can be triggered in speakers displaying a range of language proficiencies. If we combine our results with existing findings, for non-early bilinguals, the results suggest that until they have exposed to enough L2 input and become L2-dominant speakers, the effects of short-term priming are not dependent on L2 proficiency. But once they become L2-dominant, high proficiency speakers would show less priming effect as the result of suppressing one language when speaking the other.

References

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics, 38*(4), 640-653.

Bell, A. (1984), Language style as audience design. *Language in Society 13*(2): 145-204.

Boersma, P. and Weenink, D. (2011), Praat: doing phonetics by computer (Version 5.1.43). [Computer program]. Retrieved 18 July 2011, from http://www.praat.org/.

Boudahmane, K., Manta, M., Antoine., Galliano, S. and Barras, C. (2011), Transcriber: a tool for segmenting, labelling and transcribing speech (Version 1.5.1). [Computer program]. Retrieved 18 July 2011, from http://trans.sourceforge.net/en/presentation.php/.

Bradlow, A., Torretta, G. and Pisoni, D. (1996), Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication 20*(3): 255-272.

Brown, C. (1998), The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research 14*(2): 136-193.

Chang, J. (2001), Chinese speakers, in M, Swan and B, Smith, eds, *Learner English: A Teacher's Guide to Interference and other Problems*, Cambridge, pp, 310-324.

Cook, V. (1992), Evidence for multi-competence. *Language Learning* 42: 557-591.

Cook, V. (2003), *Effects of the second language on the first.* Clevedon, UK: Multilingual Matters.

Coupland, N. (1984), Accommodation at work. *International Journal of the Society of Language* 46: 49-70.

Cummins, J. (1984), *Bilingualism and Special Education: Issues in Assessment and Pedagogy*. San Diego, CA: College Hill.

Deterding, D. and Nolan, F. (2007), Aspiration and Voicing of Chinese and English Plosives, in J. Trouvain, and W. J. Barry, eds, *Proceeding of the ICPhS XVI*, Saarbrücken, pp. 917-920.

Dickerson, L. and Dickerson,W. (1977), Interlanguage phonology: current research and future direction. The notions of simplification, interlanguages, and pidgins, and their relation to second language pedagogy, in *Actes du 5ème Colloque de Linguistique Appliquée de Neuchàtel*, Geneva, Droz, pp, 18-30.

Dickerson, W. (1977), Language variation in applied linguistics. *International Review of Applied Linguistics* 35: 43-66.

Ernestus, M. T. C. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua, 142*, 27-41.

Flege, J. (1981), The Phonological Basis of Foreign Accent: A Hypothesis. *TESOL Quarterly* 15: 443-455.

Flege, J. (1987), The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47-65.

Flege, J. (1988), The production and perception of speech sounds in a foreign languages, in H. Winitz, ed, *Human communication and Its Disorders, A Review*, Norwood, NJ: Ablex, pp. 224-240.

Flege, J. (1992a), Speech learning in a second language, in C. Ferguson, L. Menn and Stoel-Gammon, eds, *Phonological Development: Models, Research, and Implications*, Timonium, MD: York Press, 565-604.

Flege, J. (1992b), The intelligibility of English vowels spoken by British and Dutch talkers, in R. Kent, ed, *Speech Disorders: Theory, Measurement, and Management,* Amsterdam, John Benjamins, pp, 157-232.

Flege, J. (1995), Second-language speech learning theory, findings, and problems, in W. Strange, ed, *Speech Perception and Linguistic Experience*, Timonium, MD: York Press, 233-277.

Flege, J. (2002), Interaction between the native and second language phonetic systems, in P. Burmeister, T. Piske and A. Rohde, eds, *An Integrated View of Language Development: Papers in Honor of Henning Wode*, Wissenschaftlicher Verlag Trier, Trier.

Flege, J., MacKay, I. and Meador, D. (1999a), Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America* 106: 2973-2987

Flege, J., MacKay, I and Piske, T. (2002), Assessing bilingual dominance. *Applied Psycholinguistics* 23: 567-598.

Flege, J., Munro, M. and MacKay, I. (1995), The effect of age of second language learning on the production of English consonants. *Speech Communication* 16: 1-26.

Flege, J. and Liu, S. (2001), The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* 23: 527-552.

Flege, J., Schirru, C., and MacKay, I. R. A. (2003), Interaction between the native and second language phonetic subsystems. *Speech Communication* 40: 467-491.

Flege, J., Yeni-Komshian, G. and Liu, H. (1999b), Age constraints on second language acquisition. *Journal of Memory and Language* 41: 78-104.

Gick, B., Wilson, I., Koch, K., & Cook, C. (2005). Language-specific articulatory settings: Evidence from inter-utterance rest position. *Phonetica, 61*(4), 220-233.

Goldinger, S.D. (2007). A complementary-systems approach to abstract and episodic speech

perception. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken. pp. 49-54.

Gordon, M., Barthmaier, P. and Sands, K. (2002), A cross-linguistic study of voiceless fricatives. *Journal of the International Phonetic Association* 32: 141-174.

Grosjean, F. (2001), The bilingual's language modes, in J. Nicol, ed, *One Mind, Two Languages: Bilingual Langauge Prociessing*, Oxford, Blackwell, pp. 1-22.

Grosjean, F. (2008). *Studying bilinguals.* Oxford: Oxford University Press.

Grosjean, F. (2012). An attempt to isolate, and then differentiate, transfer and interference. *International Journal of Bilingualism, 16*(1), 11-21.

Grosjean, F. and Miller, J. (1994), Going in and out of languages: An example of bilingual flexibility. *Psychological Science* 5: 201-209.

Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics, 31*(3), 373-405.

Hay, J., Drager, K. and Warren, P. (2009), Careful who you talk to: An effect of experimenter identity on the production of the NEAR/SQUARE merger in New Zealand English. *Australian Journal of Linguistics* 29: 269-285.

Hay, J., Drager, K. and Warren, P. (2010), Short-term Exposure to One Dialect Affects Processing of Another. *Language and Speech* 53: 447-471.

Halle, M. and Stevens, K. (1997), The postalveolar fricatives of Polish, in S. Kiritani., H. Hirose and H. Fujisake, eds, *Speech Production and Language: In Honor of Osamu Fujimura*, Berlin, Mouton & Gruyter, pp, 177-193.

Hay, J., Maclagan, M. and Gordon, E. (2008), *New Zealand English*, Edinburgh University Press.

Hazan, V. and Boulakia, G. (1993), Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech* 36: 17-38.

Ho, D. (1996), Stages and strata in dialectal history -Case studys of Heng County, Da County and Shipo, in C. Huang and Y. Li, eds, *New Horizons in Chinese Linguistics*, Dordecht, Boston, London, Kluwer, pp, 215-234.

Ho, L. (2003), Pronunciation problems of PRC students, in L. Ling, L. Ho, L. Meyer, C. Varaprasad and C. Young, eds, *Teaching English to Students from China*, Singapore, University of Singapore Press, pp, 138-155.

Holton, G. (2001), Fortis and Lenis Fricatives in Tanacross Athapaskan. *International Journal of American Linguistics* 67: 396-414.

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In *Talker variability in speech processing,* edited by Keith Johnson and John W. Mullennix, 145–165. San Diego: Academic Press.

Kang, K. and Guion, S. (2006), Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *Journal of the Acoustical Society of America* 119: 1672-1683.

Kecskes, I. and Papp, T. (2000), *Foreign language and mother tongue.* Mahwah, NJ: Erlbaum.

Kuhl, P., Tsao. F and Liu, H. (2003), Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America* 100: 9096-9101.

Kouritzin, S. (1999). *Face[t]s of first language loss*. Mahwah, NJ: Erlbaum.

Lambert, W. (1990), Persistent issues in bilingualism, in B. Harley, P. Allen, J. Cummins and M. Swain, eds, *The development of second language proficiency*, Cambridge, Cambridge University Press, pp. 201-218.

Lee, Wai-Sum & Eric Zee. (2003), Standard Chinese (Beijing). *Journal of the International Phonetic Association* 33/1: 109–112.

Lenneberg, E. (1969), *Biological foundations of language*. New York: Wiley.

Li, F., Edwards, J. and Beckman, M. (2007), Spectral measure for sibilant fricatives English, Japanese, and Mandarin Chinese, in J. Trouvain, and W. J. Barry, eds, *Proceeding of the ICPhS XVI*, Saarbrücken, pp. 917-920.

MacKay, I., Meador, D. and Flege, J. (2001), The identification of English consonants by native speakers of Italian. *Phonetica* 58: 103-125.

Major, R. (1992), Losing English as a first language. *Modern Language Journal*, 76: 190-208.

Maye, J. (2007). Learning to overcome L1 phonological biases. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken. pp. 63-66.

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 539-553.

Munro, M. and Derwing, T. (1994), Evaluations of foreign accent in extemporaneous and read material. *Language Testing* 11: 253-266.

Munro, M., Flege, J. and MacKay, I. (1996). The effects of age of second-language learning on the production of English vowels. *Applied Psycholinguistics* 17: 313-334.

Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics, 41*(6), 407-420.

Pavlenko, A. (2000), L2 influence on L1 in late bilingualism. *Issue in Applied Linguisitics* 11/2: 175-205.

Pierrehumbert, J. B. (2002). Word-specific phonetics. In: Gussenhoven, C., Warner, N., Rietveld, T. (Eds.), *Phonology & Phonetics [Laboratory Phonology 7]*. Mouton, Berlin, pp. 101-140.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech, 46*(2-3): 115-154.

Pierrehumbert, J. B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics, 2*(1), 33-52.

Piske, T., Flege, J., MacKay, I. and Meador, D. (2002), The production of English vowels by fluent Early and Late Italian-English Bilinguals. *Phonetica* 59: 49-71.

Rickford, J. and McNAIR-KNOX, F. (1994), Addressee- and topic- influence style shift: A quantitative sociolinguistic study, in D. Biber and E. Finegan, eds, *Sociolinguistic perspective on register*, New York, Oxford University Press, pp. 235-276.

Salvatierra, J. and Rosselli, M. (2001), The effect of bilingualism and age on inhibitory control. *International Journal of Bilingualism* 15: 26-37.

Sanchez, K., Hay, J., & Nilson, E. (2015). Contextual activation of australia can affect new zealanders' vowel productions. *Journal of Phonetics, 48*, 76-95.

Sancier, M. and Fowler, C. (1997), Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25: 412-436.

Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics, 43*, 26.

Su, I. (2001), Transfer of sentence processing strategies: A comparison of L2 learners of Chinese and English. *Applied Psycholinguistics,* 22: 83-112.

Su, I. (2010), Transfer of pragmatic competences: A bi-directional perspective. *The Modern Language Journal* 94: 87-102.

Sundara, M., Polka, L. and Baum, S. (2006), Production of coronal stops by stimultaneous bilingual adults. *Bilingualism: Language and Cognition* 9: 97-114.

Tarone, E. (1979), Interlanguage as Chameleon. *Language Learning* 29: 181-91.

Tarone, E. (1982), Systematicity and Attention in Interlanguage. *Language Learning* 32: 69-84.

Tarone, E. (1983), On the Variability of Interlanguage Systems. *Applied Linguistics* 4: 142-163.

Wenk, B. (1979). Ariiculatory setting and defossilization. Paper presented at the Colloque International sur l'Acquisition d'une Langue Étrangère: Perspectives de Recherche, Université de Paris-Vincennes.

Wilson, I., & Gick, B. (2014). Bilinguals use language-specific articulatory settings. *Journal of Speech, Language, and Hearing Research, 57*(2), 1-13.

Wilson, I., Horiguchi, N., & Gick, B. (2007). *Japanese articulatory setting: the tongue, lips and jaw*. Paper presented at UltraFest IV, New York University.

Wu, J. (2008), Comparison between Chinese and English Fricatives. *Journal of Mudanjiang Normal University* 5: 69-70.

Young, S. and Woodland, P. (1989-2011), HTK (Version 3.4.1). [Computer program]. Retrieved 18 July 2011, from http://htk.eng.cam.ac.uk/.

Zampini, M. (1994), The role of native language transfer and task formality in the acquisition of Spanish Spirantization. *Hispania* 77:470-481.

Appendix A

Tested tokens

sound pair	English tokens	Chinese tokens
/ʃ/-/ʂ/	shoebox	shufu
	shopping	shoupa
	sharpen	shache
	solution	qianshen
	shower	shetou
	sugar	dushe
	conscious	shuben
	Chanel	shafa

/tʃ/-/tʂʰ/	Chinese	chaihuo
	chewing	qingchen
	charger	chuzhong
	teacher	chabie
	nature	tingche
	cheesecake	qingchen
	chicken	chuqu
	touch	chedi
/dʒ/-/tʂ/	large	zheyan
	jar	zhurou
	jewelry	zhouma
	joke	guozhi
	fridge	zhanzhe
	dress	zuzhi
	just	zhaore
	Germans	liangzhi
/ʒ/-/r/	measure	huore
	vision	haoren
	usual	jinru
	treasure	jianren
	leisure	zhaore
	pleasure	menre
	casual	qianrang
	decision	zhurou
English /r/	sara	
	rabbit	
	really	
	rifling	
	various	
	products	
	present	
	jewellery	
START-/a/	sharpen	shafa
	jar	chabie
	charger	fazhan

	faster	paqie
	passing	zhaying
	large	
GOOSE-/u/	shoebox	shufu
	usual	chuqu
	sugar	zhurou
	chewing	chuzhong
	bootee	bushi
	jewellery	
FLEECE-/i/	cheesecake	caimi
	meatball	pifu
	keeping	tichu
	weekend	mihu
	teaspoon	tizi
	eaten	yige
DRESS-/e/	better	yewan
	dressing	yeye
	bedroom	yezi
	metal	yewai
	getting	yeti
TRAP-/e/	apple	yewan
	happen	yeye
	planet	yezi
	rabbit	yewai
	blankets	yeti
LOT-/o/	shopping	boli
	lotto	boxue
	hotter	bozi
	bottom	moshui
		podu
THOUGHT-/o/	thoughtful	boli
	naughty	boxue
	sorter	bozi
	shorter	moshui
		podu

Appendix B

Reading material in each language

English

Sara woke to the sound of the howling wind. She lay in her bedroom, barely conscious, knowing that she would probably have to give up her weekend plans; she snuggled under the blankets, keeping warm. Sara was a teacher. She also worked on Sundays as a mail sorter. Today was her one day off, and she had been planning to buy a pair of bootees for one of her colleagues, who had a new baby. She had also hoped to go out to lunch, but now this seemed unlikely to happen. Drifting to the kitchen in her dressing gown, she opened the fridge. It only contained some leftover pizza, some meatballs and a cheesecake with some mould on the bottom. Yuck! She didn't really want to touch these: She made a coffee, and took a large jar of sugar off the shelf. She measured out two teaspoons into her coffee, and put the cup into the microwave to make it a bit hotter. Then she fed the pizza to the cats, and the cheesecake to the rabbits. She liked having pets, despite the fact that the rabbits could be a bit naughty sometimes. She waited for the rain to

clear, but it didn't. She finally made the decision to go out anyway. It was no joke having no decent food in the house. Going out was the only solution.

Fortunately, her house was just five-minutes walking distance from the closest mall. After taking a shower and getting dressed casually, she walked through the rain, passing the park and heading to the mall. She walked faster than usual, listening to heavy metal on her iPod. The mall was a big mall, with lots of shops, a movie theatre, and a leisure centre. She liked shopping, and took great pleasure in rifling amongst the various products, and seeking out a perfect pair of bootees for the gift. She finally chose some and had them wrapped in a tiny shoebox. She felt satisfied with the sweet, thoughtful present. She bought a charger for her new phone. She tried on a dress, but it was too long, and would need to be made considerably shorter, so she decided not to buy it. Then she stopped at the Chanel shop, and checked out their hats, boots and shoes. She also lingered at the jewellery display window, containing many trinkets and treasures. She had a vision of being able to wear the jewellery and so decided to buy a lotto ticket. Finally, she went into a Chinese restaurant and ordered some chicken. The chef sharpened the knife and expertly chopped the meat into even cuts, Sara sat, chewing happily. She instantly felt better. The TV in the restaurant was tuned to the Nature channel, but, strangely it was playing a football game: The Germans against the Spanish. After being asked to change it, a waiter changed it to the Discovery Channel, which was introducing planets in the solar system. After she had eaten, she walked home, past the fruit market, where she bought some apples and bananas for an afternoon snack.

Chinese

和许多人一样，很多年前爷爷带着老伴张奶奶把家从城里搬到了城外，他们只有一个女儿，虽然她并不漂亮做人也很迷糊，但是却很善良，也懂得谦让，而且有坚韧的性格，大家都说她是个好人。当女孩上初中的时候，爷爷给家里添了一台二手车，方便接送她，不可否认，这些年经济发展了很多，他总是把车停在院子里。每天清晨，奶奶会往灶炉里填满柴火，煎好猪肉饼，在他们各自的杯子里盛上果汁，等爷俩起身。

通常爷爷会第二个起身，伸展下身体，便出去先把停在屋外的车子玻璃擦干净，并彻底检查一次发动机和刹车是否一切正常，然后回到屋里，在沙发上舒服地坐下，拿起放在旁边

的书本，爷爷很博学，昨天放下时正好看到诗词的部分。待奶奶做好早餐之后爷爷会上楼催促女儿起身，并督促她检查书包是否收好，并将放在衣橱里的手帕和上书法课要用到的墨水塞到她的书包里，女孩这时总是会调皮的吐下舌头。女孩一直粗心，总是丢三落四的。
这天，吃早餐的时候，女孩提出说学校组织去野外的春游，需要奶奶为她准备一些炒菜和液体状的流食，因为夜晚需要在外扎营，老师让大家带上锤子等工具。爷爷听到后说，家里的锤子上次被送修屋顶的梯子的少年扒窃了，等送她去上学后，他会去再买一把，说着伸手挠了挠脖子，皮肤马上红了一片。紧接着，站着清洗灶台的奶奶提醒女孩，去山里不要凑近落在灌木丛中的落叶，以免招惹到用叶子遮掩藏的毒蛇，说完转身回屋里拿出前几天一直在做的刺绣，说是要赶在女孩春游前弄好，好让她带着漂亮的去玩。
初夏天，一过中午气温就变得火热，闷热难当。然而即便如此，女孩和她认识的同学还是兴致勃勃的朝扎营地儿进发。走着走着，眼前出现一个坡度很大的大山包，山体的下方有个先人挖好的隧道，里面很黑，在大家即将进入的时候，有两只松树从洞里窜了出来，不知道是谁因为受到惊吓咒骂的一声。
隧道不长，可以看到对面透过来的光，不是那种完全漆黑的山洞。出了隧道，当大家看到在不远处的那条清澈的小溪水都变得有些兴奋。溪边的枯木上长满了木耳，有的同学浴室采摘了一大包。女孩卸下背包跑到小溪旁欠身用冰凉的溪水将自己的脸和双耳都清洗了一遍，顿时觉得清爽了许多
以前不是没有来扎过营，这次的差别在于不是和父母，而是和同学一块，晚上大家可以围着篝火做猜谜游戏，想想就觉得高兴。希望下次还有这样的机会。

Appendix C

Proficiency scores for each speaker

	% of Chinese	friend	AOA	LOR	partner	working	degree	total
speaker3	4	0	-7	7	0	3	2	9
speaker4	6	0	-8	7	1	3	1	10
speaker5	8	3	-8	8	1	0	1	13
speaker7	8	3	-6	6	1	3	2	17
speaker8	4	0	-6	6	0	3	1	8
speaker12	4	0	-8	9	0	0	3	8
speaker13	4	0	-7	6	1	3	2	9
speaker14	6	0	-12	10	0	3	4	11
speaker1	4	0	-7	2	1	0	2	2

speaker2	2	0	-4	8	0	0	1	7
speaker6	4	0	-7	8	0	0	2	7
speaker9	0	0	-13	2	0	0	3	-8
speaker10	4	0	-11	5	1	0	3	2
speaker11	0	0	-7	4	1	3	2	3

Appendix D

Questionnaires

Part1
#_____

BACKGROUND QUESTIONNAIRE ENGLISH

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Thank you for taking part in our study.
Before you read the passages, we have a few questions, which will help us with our analysis.

1. Age:_____

2. Sex: M F

3. Which part of China do you come from? _____

4. Which part of China do your parents come from (if different from (1)).

   _____

5. Where do you live now?_____

6. Where do your parents live now? _____

7. How old were you when you came to New Zealand? _____

8. Do you intend to stay in New Zealand permanently? _____

9. What is your occupation (i.e. student, dentist, etc.)?_____

10. What's the highest educational qualification you have received?

    _____

11. Do you speak any languages apart from English and Chinese? If so, please list them here:

12. What is/was your Mother's occupation?_____

13. What is/was your Father's occupation? _____

Part2

BACKGROUND QUESTIONNAIRE MANDARIN (to be translated)

1. What dialect of Chinese is your native dialect?

2. What is the ethnicity of your partner if you have one _____

3. How often do you go to China? _____

4. What language do you normally speak at home? _____ (if Chinese, please be specific about dialect)

5. If you work, what language do you normally speak at work? (if Chinese, please be specific about dialect)_____

6. If you attend university, what language do you normally speak at university? (if Chinese, please be specific about dialect)_____

7. During your interactions in a typical day, what proportion of the time would you spend speaking standard Mandarin?

   0%20% 40% 60% 80% 100%

8. During your interactions in a typical day, what proportion of the time would you spend speaking another dialect of Chinese?
9.

   0% 20% 40%60%80%100%

(please indicate which dialect here: _____)

10. Would you estimate amongst your close friends there are:

(a) More Native speakers of English
(b) More Native speakers of Chinese
(c) Roughly equal numbers of speakers of Chinese and English

11. Do you have children?

If so:

- How old are your children?
- What language do you usually speak to your children in?
- What language do your children usually speak to you in?

(Note that the experiment only used translated version for part2, thus speakers only saw the Chinese version)

Translation of part2 (in simplified Chinese)
#_____

汉语方言背景问卷

1. 你说那种汉语方言（如果你说的话）？_____

2. 如果您有配偶的话，他/她的种族是什么？_____

3. 您多久去一次中国？_____

4. 您通常在家说什么语言？（如果是汉语，请写出哪种方言）_____

5. 如果您有工作，你在工作时通常说什么语言？（如果是汉语，请写出哪种方言）_____

6. 如果您在上大学，您在大学里通常说什么语言？ （如果是汉语，请写出哪种方言）_____

7. 在您一天与他人的交流中， 您说普通话的比例大概有多少？
   0% 20% 40% 60%80% 100%
8. 在您一天与他人的交流中， 您说汉语方言的比例大概有多少？

   0% 20% 40% 60%80% 100%

   (请写出方言种类_____)
9. 在您的关系好的朋友中，你是否可以估计一下他们

   （a） 大部分人说英语为母语的
   （b） 大部分人说汉语为母语的
   （c） 他们的数量大概相等
10. 您有孩子吗？
    如果有：
    他们几岁了_____
    您通常用那种语言（或方言）和他们交谈？ _____
    您的孩子通常用那种语言（或方言）和您交谈？_____

Appendix E

Information Sheet

Research Information
Full Project Title: Bilingual Speech Production Experiment

- **Your Consent**

You are asked to participate in a research study conducted by Keyi Sun as the main part of the course requirements for a Masters Thesis at the University of Canterbury. This work is conducted under the supervision of Associate Professor Jennifer Hay. We are interested in how bilingual speakers pronounce certain vowels and consonants. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

- **PARTICIPATION AND WITHDRAWAL**

Your participation in this study is completely voluntary and you are free to choose whether to be in it or not. If you choose to be in this study, you may subsequently withdraw from it at any time without penalty or consequences of any kind. Withdrawal of participation includes withdrawal of any information already provided. The investigator may withdraw you from this research if circumstances arise which warrant doing so. For example, such circumstances include the possibility that your background is not compatible with the desired demographic for this study, or that the recording equipment may be misfunctioning during the interview

- **PURPOSE OF THE STUDY**

The aim of this project is to investigate pronunciation patterns in bilingualism

- **PROCEDURES**

Your involvement in this project will involve reading two passages: one in English and one in Chinese. You can ask to have either simplified or traditional Chinese. You will also be asked to return to the second part of the experiment at least a week after the first part. Your voice will be recorded in the experiment.

- **POTENTIAL RISKS AND DISCOMFORTS**

There are no potential risks or discomforts.

- **POTENTIAL BENEFITS**

This research may have implications for models of bilingualism, and theories of how multiple languages are stored in the mind
There will be no benefits to you, personally. If you would like a copy of any published paper reporting the results, please email ksu25@student.canterbury.ac.nz

- **PAYMENT FOR PARTICIPATION**

Participation in this experiment is voluntary

- **CONFIDENTIALITY**

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law.

The results of this study may be published, but your anonymity will be preserved. You will be identified by number, not by name.

•IDENTIFICATION OF INVESTIGATORS

If you have any questions or concerns about the research, please feel free to contact Keyi Sun at ksu25@uclive.ac.nz or the supervisor Jennifer Hay at jen.hay@canterbury.ac.nz. They would be pleased to discuss any concerns you may have about participation in the project.

•RIGHTS OF RESEARCH SUBJECTS

You are not waiving any legal claims, rights or remedies because of your participation in this study.

•HUMAN ETHICS COMMITTEE APPROVAL

This study has been reviewed and approved by the Human Ethics Committee at the University of Canterbury.

Appendix F

Consent Form

Consent Form

Full Project Title: Bilingual Speech Production Experiment

HUMAN ETHICS COMMITTEE APPROVAL

This study has been reviewed and approved by the Human Ethics Committee at the University of Canterbury.

I have read and I understand the procedures described in the Research Information.
My questions have been answered to my satisfaction, and I agree to participate in this study.

Name of Subject _____

Signature _____ Date _____

Appendix G

Debriefing Sheet

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Debriefing sheet

Bilingual Speech Production Experiment

Thank you for participating in this study. We will be analyzing your pronunciation in the two reading passages, with particular attention to pronunciation of vowels, and sounds like ʃ and ʒ in English, and sh and r in Chinese. We are interested in whether pronunciation of these sounds in English and Chinese is affected by :

1. How long you have been learning English, and

2. Whether you have just been speaking English or Chinese.

Thank you for your help.